

**RECONOCIMIENTO DE ACTIVIDAD FÍSICA HUMANA USANDO
MÁQUINAS DE SOPORTE VECTORIAL PENALIZADAS Y MODELOS
OCULTOS DE MARKOV EN SISTEMAS MULTIMODALES**

LEIDY ESPERANZA PAMPLONA BERON

Proyecto de grado presentado como requisito parcial
para aspirar al título de Magister en Ingeniería Eléctrica

Director

Ing. Andrés Felipe Calvo Salcedo, M.Sc.

**UNIVERSIDAD TECNOLÓGICA DE PEREIRA
FACULTAD DE INGENIERÍAS
PROGRAMA DE MAESTRIA EN INGENIERÍA ELÉCTRICA
LINEA DE AUTOMÁTICA
PEREIRA
2019**

Nota de Aceptación

Firma del Presidente del jurado

Firma del jurado 1 - Evaluador

Firma del jurado 2 - Evaluador

Firma del jurado 3 - Director

Pereira, 04 de Febrero de 2019

Agradecimientos

Inicialmente quiero agradecerle a Dios por permitirme llegar a este momento de finalización de un ciclo importante en mi formación profesional, al cual dedique un tiempo significativo en el que mi madre *Ofelia* y mis hermanas *Elizabeth* y *Alejandra* me tuvieron paciencia y sabiduría para manejarlo y apoyarme en todo los momentos. Gracias *Julián* por ayudarme en aquellas tareas que aportaron en el desarrollo de este trabajo. A mi familia donde siempre hubo una sonrisa y palabra de aliento para continuar.

Agradezco a mi tutor de la maestría Andrés Felipe Calvo, que sin pensar me mostró este camino y dio un voto de confianza para poder llevarlo a cabo, a Carlos por su tiempo y asesoría, Daniel Molina por brindarme su amistad , Arley Bejarano por compartir esos momentos de frustración pero también de felicidad que se tuvieron en este trayecto, a mis compañeros de GIIE Andrés Castillo y Cristian Escudero por su colaboración en este tiempo.

También agradecerle al programa de Ingeniería Eléctrica e Ingeniería Electrónica y a la Universidad Tecnológica de Pereira por su apoyo durante la maestría y darme la oportunidad de avanzar un escalón más en mi vida profesional, a Clarita y Beatriz por animarme y escucharme y para aquellas personas que participaron directa o indirectamente en este ciclo.

Pocos son los amigos, pero más importante es poder contar con ellos en todo momento en este camino que no esta escrito.

Muchas Gracias.

CONTENIDO

pág.

TÍTULO

CONTENIDO

| | |
|--|-----------|
| 1. INTRODUCCIÓN | 11 |
| 1.1. DEFINICIÓN DEL PROBLEMA | 12 |
| 1.2. JUSTIFICACIÓN | 15 |
| 1.3. OBJETIVOS | 16 |
| 1.3.1. Objetivo General | 16 |
| 1.3.2. Objetivos Específicos | 16 |
| 2. ESTADO DEL ARTE | 17 |
| 3. MARCO CONCEPTUAL | 21 |
| 3.1. Unidad de Masa Inercial (IMU) | 21 |
| 3.2. Electromiografía (EMG) | 21 |
| 3.3. Kinect | 21 |
| 3.4. Movimientos primitivos | 21 |
| 3.5. Fusión sensorial | 22 |
| 3.6. Ventana | 22 |
| 3.7. Análisis de componentes principales (PCA) | 23 |
| 3.8. Extracción de características | 23 |
| 3.9. Clasificación | 23 |
| 3.9.1. Clasificación Supervisada | 23 |
| 3.9.2. Clasificación No Supervisada | 24 |
| 3.10. Análisis de Desempeño de un clasificador | 24 |

| | |
|---|-----------|
| 4. MARCO TEÓRICO | 27 |
| 4.1. Máquinas de vectores de soporte (SVM) [1, 2] | 27 |
| 4.1.1. Funciones Kernel [3] | 29 |
| 4.1.2. SVM Multiclase [4] | 30 |
| 4.1.3. Método 1 de penalización: C-Clasificación de los Vectores de Soporte(C-SVC) [5] | 31 |
| 4.1.4. Método 2 de penalización: SVM binaria pondera- da [6] | 32 |
| 4.2. Modelos Ocultos de Markov (HMM) [7] | 34 |
| 5. CAPTURA DE DATOS DEL SISTEMA MULTIMODAL | 37 |
| 5.1. Base de datos anotada de actividades físicas humanas . . | 39 |
| 5.2. Base de datos anotada de movimientos primitivos | 43 |
| 6. METODOLOGÍA | 49 |
| 7. RECONOCIMIENTO DE MOVIMIENTOS PRIMITIVOS | 51 |
| 7.1. Codificación de movimientos primitivos | 51 |
| 7.1.1. Tamaños de ventana | 53 |
| 7.1.2. Extracción de características del Kinect One® . . | 54 |
| 7.1.3. Extracción de características para red sensores IMU | 56 |
| 7.1.4. Extracción de características Red de sensores EMG | 59 |
| 7.2. Clasificación de movimientos primitivos | 59 |
| 8. RECONOCIMIENTO DE ACTIVIDADES | 61 |
| 8.1. Fusión de datos | 61 |
| 8.2. Clasificación, entrenamiento y validación del modelo . . . | 61 |
| 9. EXPERIMENTOS Y RESULTADOS | 63 |
| 9.1. Tamaños de ventanas | 65 |
| 9.2. Análisis de movimientos primitivos | 66 |

| | |
|---|----|
| 9.3. Reconocimiento de actividades físicas | 69 |
| 9.3.1. Experimento Kinect One®+IMUs | 71 |
| 9.3.2. Experimento Kinect One®+EMGs | 71 |
| 9.3.3. Experimento IMUs+EMGs | 72 |
| 9.3.4. Experimento Kinect One®+IMUs+EMGs | 73 |
| 9.3.5. Comparación del desempeño del clasificador con todo el conjunto de sensores para cada modelo de la sección 9.2 | 73 |
| 10. CONCLUSIONES Y RECOMENDACIONES | 77 |
| 10.1. Conclusiones | 77 |
| 10.2. Recomendaciones | 78 |
| BIBLIOGRAFÍA | 81 |
| 11. ANEXOS | 89 |
| 11.1. Matrices de Confusión de los movimientos primitivos . . . | 89 |
| 11.1.1. Resultados experimentales con la metodología de SVM_CyH | 89 |
| 11.1.2. Resultados experimentales con la metodología de SVM_C | 90 |
| 11.1.3. Resultados experimentales con la metodología de SVM_BP | 91 |
| 11.1.4. Resultados experimentales con la metodología de C-SVM_V_A-Corr | 92 |
| 11.2. Matrices de Confusión de las actividades físicas | 94 |
| 11.2.1. Resultados experimentales con la metodología de HMM_CyH | 94 |
| 11.2.2. Resultados experimentales con la metodología de HMM_C | 96 |

| | |
|---|----|
| 11.2.3. Resultados experimentales con la metodología de | |
| HMM_BP | 98 |

LISTA DE TABLAS

| | | |
|-----|--|----|
| 1. | Lista de actividades. | 42 |
| 2. | Lista de movimientos primitivos. | 44 |
| 3. | Etiquetas de la SVM para cada postura | 52 |
| 4. | Descriptores estadísticos utilizados. | 55 |
| 5. | Puntos articulados del Kinect One®. | 56 |
| 6. | Características IMU. | 58 |
| 7. | Descriptores estadísticos utilizados | 58 |
| 8. | Tamaños de ventanas. | 66 |
| 9. | Traza de las matrices de confusión con Kinect One®. | 67 |
| 10. | Diagonal de las matrices de confusión con la red de sensores IMU. | 67 |
| 11. | Diagonal de las matrices de confusión con la red de sensores EMG. | 68 |
| 12. | Comportamiento promedio de la metodología para el reconocimiento de los movimientos primitivos. | 69 |
| 13. | Matriz de confusión para el reconocimiento de actividades físicas con el Kinect One®. | 70 |
| 14. | Matriz de confusión para el reconocimiento de actividades físicas con IMUs. | 70 |
| 15. | Matriz de confusión para el reconocimiento de actividades físicas con EMG. | 70 |
| 16. | Matriz de confusión Kinect One®+IMUs. | 71 |
| 17. | Matriz de confusión Kinect One®+EMGs. | 72 |
| 18. | Matriz de confusión IMUs+EMGs. | 72 |
| 19. | Matriz de confusión Kinect One®+IMUs+EMGs. | 73 |
| 20. | Comparación del desempeño con diferentes sensores y combinaciones de estos con el modelo HMM_CyH. | 74 |
| 21. | Comparación del desempeño con diferentes sensores y combinaciones de estos con el modelo HMM_C. | 74 |

| | | |
|-----|---|----|
| 22. | Comparación del desempeño con diferentes sensores y combinaciones de estos con el modelo HMM_BP. | 75 |
| 23. | Comportamiento promedio de la metodología con una modalidad de sensor para cada modelo de clasificación. | 75 |
| 24. | Comportamiento promedio de la metodología con la fusión de sen- sores para cada modelo de clasificación. | 75 |

LISTA DE FIGURAS

| | | |
|-----|--|----|
| 1. | Esquema de la distribución de los sensores y del sistema de captura sincronizado. | 39 |
| 2. | Adquisición del dispositivo Kinect One®. | 41 |
| 3. | Adquisición de la red de sensores EMG. | 41 |
| 4. | Adquisición de la red de sensores IMU. | 41 |
| 5. | Diseño de la interfaz gráfica de captura de actividades. | 42 |
| 6. | Estructura de la base de datos. | 43 |
| 7. | Ángulo de flexión y extensión de la rodilla. a) Pierna izquierda, b) Pierna derecha. | 46 |
| 8. | Umbrales para la asignación de etiquetas de los movimientos primitivos. | 48 |
| 9. | Metodología para el reconocimiento de actividades físicas humanas. | 50 |
| 10. | Actividades 1 y 3. | 52 |
| 11. | Actividades 2 y 8. | 52 |
| 12. | Actividades 4, 6 y 7. | 53 |
| 13. | Actividad 5. | 53 |
| 14. | Actividades 9 y 10. | 53 |
| 15. | Matriz de confusión Movimientos primitivos con Kinect One®. | 89 |
| 16. | Matriz de confusión Movimientos primitivos con IMUs. | 89 |
| 17. | Matriz de confusión Movimientos primitivos con EMGs. | 89 |
| 18. | Matriz de confusión Movimientos primitivos con Kinect One®. | 90 |
| 19. | Matriz de confusión Movimientos primitivos con IMUs. | 90 |
| 20. | Matriz de confusión Movimientos primitivos con EMGs. | 91 |
| 21. | Matriz de confusión Movimientos primitivos con Kinect One®. | 91 |
| 22. | Matriz de confusión Movimientos primitivos con IMUs. | 91 |
| 23. | Matriz de confusión Movimientos primitivos con EMGs. | 92 |
| 24. | Matriz de confusión Movimientos primitivos con Kinect One®. | 92 |

| | | |
|-----|---|-----|
| 25. | Matriz de confusión Movimientos primitivos con IMUs. | 93 |
| 26. | Matriz de confusión Movimientos primitivos con EMGs. | 93 |
| 27. | Matriz de confusión de actividades físicas con Kinect One®. | 94 |
| 28. | Matriz de confusión de actividades físicas con IMUs. | 94 |
| 29. | Matriz de confusión de actividades físicas con EMGs. | 94 |
| 30. | Matriz de confusión de actividades físicas con Kinect One®+IMUs. | 95 |
| 31. | Matriz de confusión de actividades físicas con Kinect One®+EMGs. | 95 |
| 32. | Matriz de confusión de actividades físicas con IMUs+EMGs. | 95 |
| 33. | Matriz de confusión de actividades físicas Kinect One®+IMUs+EMGs. | 96 |
| 34. | Matriz de confusión de actividades físicas con Kinect One®. | 96 |
| 35. | Matriz de confusión de actividades físicas con IMUs. | 96 |
| 36. | Matriz de confusión de actividades físicas con EMGs. | 97 |
| 37. | Matriz de confusión de actividades físicas con Kinect One®+IMUs. | 97 |
| 38. | Matriz de confusión de actividades físicas con Kinect One®+EMGs. | 97 |
| 39. | Matriz de confusión de actividades físicas con IMUs+EMGs. | 98 |
| 40. | Matriz de confusión de actividades físicas Kinect One®+IMUs+EMGs. | 98 |
| 41. | Matriz de confusión de actividades físicas con Kinect One®. | 98 |
| 42. | Matriz de confusión de actividades físicas con IMUs. | 99 |
| 43. | Matriz de confusión de actividades físicas con EMGs. | 99 |
| 44. | Matriz de confusión de actividades físicas con Kinect One®+IMUs. | 99 |
| 45. | Matriz de confusión de actividades físicas con Kinect One®+EMGs. | 100 |
| 46. | Matriz de confusión de actividades físicas con IMUs+EMGs. | 100 |
| 47. | Matriz de confusión de actividades físicas Kinect One®+IMUs+EMGs. | 100 |

1. INTRODUCCIÓN

El reconocimiento de las actividades físicas se estudia actualmente con el objetivo de comprender las acciones de las personas y de como estas interactúan con su entorno físico. Algunas de las áreas aplicaciones son los vídeo juegos, robótica, rehabilitación, ingeniería deportiva, seguridad, entre otras [8, 9, 10, 11]. Por lo tanto, para la identificación de las actividades se debe realizar el seguimiento de los movimientos efectuados por el cuerpo humano, considerando que este se modela como un conjunto de articulaciones que interactúan entre sí para efectuar una acción [8, 11, 12].

Para la detección de las actividades físicas se han usado diferentes modalidades de sensores tales como cámaras de profundidad (Kinect®, MoCap), cámaras de vídeo, sensores ambulatorios (IMUs) y sensores electromiográficos (EMGs). Cada sensor tiene ventajas significativas con respecto a otras modalidades para la caracterización de actividades, donde la fusión de estos permite mejorar la eficiencia de los clasificadores en el reconocimiento de las actividades. Entre los trabajos que han desarrollado una metodología de fusión se tienen las siguientes combinaciones Kinect®+IMU, Kinect®+EMG y EMG+IMU. Sin embargo, sólo en [13] y [14] realizan la fusión de las tres modalidades de sensores mencionadas con anterioridad obteniendo mejores rendimientos con respecto a los métodos que usan uno o dos tipos de sensores [15, 16, 17, 18, 19].

Una actividad se puede representar como una secuencia continua de acciones conocidas como movimientos primitivos, que es la segmentación del movimiento en componentes más pequeñas configuradas en un orden específico para cada actividad. Para la etiquetación de estos sub-movimientos se han utilizado técnicas supervisadas y no supervisadas, ya sea de forma manual o por algoritmos de clustering, K-means o modelos ocultos de Markov (HMM) [14, 20, 21]. Entre las características que se extraen para el clasificador se tienen métricas estadísticas, espectrales, de distancia, ángulos y aceleración de los puntos articulados estableciendo una ventana de tamaño fijo o dinámico [21, 22].

Una de las técnicas de aprendizaje de máquina más utilizadas para el reconocimiento de actividades físicas es la combinación de las máquinas de vectores de soporte (SVM) y los mo-

delos oculto de Markov (HMM). Las SVM permiten clasificar los datos de forma individual o en pequeñas ventanas de tiempo para generar las etiquetas de los movimientos primitivos, mientras que los HMM encuentran la secuencia más probable de primitivos para la identificación de la actividad ejecutada [13, 20, 23, 24]. En [13] se desarrolla esta metodología para la identificación de 5 actividades fusionando tres tipos de sensores logrando desempeños superiores al 95 % para cada actividad. De acuerdo a lo anterior, el aporte de este trabajo consiste en extender la base de datos de [13] en 10 actividades físicas con el registro sincronizado de tres modalidades de sensor (Kinect One®, IMUs y EMGs), junto con el estudio comparativo de diferentes modelos de aprendizaje de máquina para la identificación de los movimientos primitivos. Para el reconocimiento de las actividades se utilizarán los HMM que permitirán la clasificación de las secuencias de movimientos correspondientes a cada acción. Esta metodología sigue el trabajo propuesto por *Andrés et al* [13], ya que es el único que realiza la fusión de estas tres modalidades de sensores usando las máquinas de soporte vectorial y los modelos ocultos de Markov para la clasificación de los sub-movimientos y actividades respectivamente.

1.1. DEFINICIÓN DEL PROBLEMA

Diferentes estudios enfocados al reconocimiento de actividades físicas humanas utilizan varias modalidades de sensores para la captura de movimiento. Estos sensores presentan limitaciones para el registro de ciertas actividades de acuerdo a su configuración y cantidad. Algunos de los problemas que se evidencian durante el proceso de adquisición son la pérdida de información, el rango de cobertura espacial, la oclusión, la precisión e incertidumbre [10]. Entre los sensores que se suelen utilizar para el reconocimiento de actividades son el Kinect®, IMU, EMG, entre otros [9, 10]. Es importante definir la frecuencia en que se va a muestrear cada sensor para no perder detalles de las actividades al momento de su análisis [9], sin embargo, una alta tasa de muestreo no implica una mejor precisión, ya que muchos de los datos adquiridos pueden contener ruido [8].

El Kinect® es un dispositivo que permite capturar el movimiento de forma no invasiva. Este dispositivo registra diferentes tipos de datos tales como los puntos articulados del esqueleto

en 3D, mapas de profundidad y vídeo. No obstante, presenta algunos problemas de oclusiones parciales del objetivo generando secuencias de ruido y haciéndolo menos preciso [14, 1, 25, 26]. En otros trabajos se ha efectuado la fusión de diferentes tipos de sensores de unidad de masa inercial (IMU) tales como acelerómetros, giroscopios y magnetómetros mostrando eficiencias superiores al 91 % en diferentes metodologías para la clasificación de actividades [8, 1]. Así mismo, los electrodos permiten la adquisición de señales electromiográficas (EMG) midiendo la contracción y extensión de los músculos [1]. Aunque, tanto los IMUs como los EMGs requieren de más de un sensor para el registro completo de diferentes actividades [9, 1].

En [27] se han fusionado diferentes modalidades de sensores tales como IMU y Kinect® mediante un filtro de Kalman adaptativo (EKF) obteniendo un error cuadrático medio (RMSE) menor al 3 % en comparación con los resultados individuales. En [13] se realiza la fusión del Kinect®, IMU y EMG con la combinación de las SVM y los HMM alcanzando eficiencias mayores al 95 % en comparación con los resultados individuales. Por lo tanto, realizar la combinación de múltiples sensores tiene sus ventajas al momento de mejorar la precisión en la identificación de las actividades [10]. No obstante, esto genera una alta dimensionalidad en los datos a procesar con un alto costo computacional. Algunos métodos que se pueden implementar para reducir la dimensionalidad de los datos son PCA, KPCA, MDS y t-SNE determinando el conjunto mínimo de características de los datos para su respectiva clasificación [11, 14, 22].

Para la clasificación de actividades físicas se han desarrollado varias metodologías utilizando técnicas de aprendizaje de máquina como KNN, Máquinas de vectores de soporte (SVM), redes neuronales, modelos ocultos de Markov (HMM), entre otros. En algunos enfoques se realiza inicialmente la detección de los movimientos primitivos como en [14, 1, 28]. Inicialmente se define un tamaño de ventana para la segmentación los sub-movimientos que depende las variaciones que se tengan entre clases [14]. En otros casos se construye un libro de códigos con las posturas claves para cada una de las actividades. Sin embargo, esta metodología sólo puede ser aplicada con los puntos articulados del esqueleto lo que limita el uso a otras modalidades de sensor [23]. Para cada uno de estos enfoques es importante determinar la transición entre cada

uno de los sub-movimientos mediante algoritmos de agrupamiento (clustering) o de segmentación para dividir una actividad en varios grupos [8, 14, 23]. Estos métodos de segmentación dividen el tiempo de la actividad en intervalos homogéneos que no están traslapados usando herramientas tales como la correlación, media y desviación [29], velocidad de los movimientos [27], entre otros. Por otro lado, también se determina si la actividad se asocia a un movimiento natural (caminar, trotar, saltar, etc.) o a un estado estático (sentarse, levantarse, etc.) donde se considera la frecuencia de muestreo de cada sensor y un tamaño de ventana, con el fin de establecer una transición lógica para cada acción y evitar traslapes entre clases[8].

Independientemente de la modalidad de sensor usada, el reconocimiento de actividades es un problema típico de clasificación donde es evidente que la combinación de diferentes modalidades de sensores permite realizar la extracción de características significativas para el modelo de aprendizaje de máquina. No obstante, son pocos los trabajos donde se realiza la fusión de un sistema multimodal conformado por Kinect®, IMU y EMG, siendo estos sensores los más utilizados para la identificación de actividades. Por lo que este estudio parte del trabajo desarrollado en [13], donde se construyó una base de datos limitada a sólo 5 actividades altamente separables entre sí. Así mismo, el desempeño de las SVMs en la generación de etiquetas de los movimientos primitivos para cada sensor es aceptable. Donde el Kinect presenta una eficiencia del 85 % aproximadamente en comparación con las otras modalidades, esto se debe a las oclusiones parciales o auto-occlusiones del Kinect® o pérdidas de conexión en los sistemas de comunicación inalámbrica en la adquisición de los datos de los sensores IMU o EMG. Mientras que en [14] con la misma base de datos utilizando técnicas de aprendizaje no supervisado como Spectral Clustering y K-means Clustering para el análisis de los sub-movimientos se obtuvo eficiencias menores al 80 %, por lo que se plantea la necesidad de desarrollar una metodología que permita penalizar los datos que no pertenezcan a cada movimiento por medio de las máquinas de vectores de soporte, con el fin de minimizar el error de entrenamiento relajando el margen del clasificador.

1.2. JUSTIFICACIÓN

En diferentes trabajos enfocados en el reconocimiento de actividades físicas humanas han utilizado diferentes tipos de sensores como el Kinect®, IMU y EMG para el registro de las actividades [23, 1, 25, 26, 27, 28]. A su vez la fusión estos sensores (Kinect-IMU, Kinect-EMG, Kinect-IMU-EMG) permiten mejorar el desempeño del clasificador en la identificación de actividades [10, 14, 1, 27].

Independiente de la modalidad de sensor, el reconocimiento de las actividades físicas humanas se considera un problema de clasificación. La selección de características de cada sensor influyen en el desempeño del modelo de aprendizaje de máquina [11]. Sin embargo, analizar el conjunto total de características genera una alta dimensionalidad de datos aumentando el costo computacional en el proceso de entrenamiento. Métodos como Análisis de Componentes Principales (PCA) y Análisis de Componentes Independientes (ICA) permiten reducir el conjunto de características obteniendo resultados confiables en la clasificación disminuyendo su costo computacional. De esta manera se pueden implementar este tipo de metodologías en sistemas embebidos [11].

Uno de los métodos que se han trabajado para la identificación de actividades físicas consiste en realizar el análisis de los movimientos primitivos para luego identificar la secuencia más probable correspondiente a la actividad [8, 14, 23, 1, 28]. En este sentido, es importante definir un tamaño de ventana durante la segmentación del movimiento donde se considere la frecuencia de muestreo de cada sensor. [8, 9, 23, 27].

Entre las metodologías empleadas para la detección de los movimientos primitivos utilizan técnicas de aprendizaje supervisados y no supervisado. En [1] se emplearon las máquinas de vectores de soporte para la clasificación de los sub-movimientos, sin embargo, no se penalizaron los datos que no estaban asociados a cada clase. En [14] utilizan clustering, K-means y K-vecinos para la codificación de los movimientos primitivos donde la eficiencia del clasificador varía dependiendo de la cantidad de clusters.

De acuerdo a lo anterior, para el reconocimiento de actividades físicas con sistemas multimodales (Kinect, IMU y EMG) es fundamental desarrollar una metodología que permita

realizar una correcta identificación de los movimientos primitivos, seleccionando un tamaño de ventana y penalizando aquellos datos que sean atípicos o que estén muy cercanos al margen de la máquina de vectores de soporte, esto con el fin de mejorar la eficiencia del clasificador al momento de reconocer cada una de las actividades.

1.3. OBJETIVOS

1.3.1. Objetivo General

Desarrollar una metodología para la clasificación automática de actividades físicas humanas utilizando penalización de datos atípicos.

1.3.2. Objetivos Específicos

1. Crear una base de datos multimodal de diez actividades físicas humanas.
2. Determinar el tamaño de ventana para describir un movimiento primitivo.
3. Determinar una metodología que penalice los vectores de soporte en las SVM para la clasificación de los movimientos primitivos.
4. Implementar la metodología para el reconocimiento de las actividades.
5. Validar estadísticamente la metodología diseñada y realizar la comparación de otras metodologías implementadas en la base de datos.

2. ESTADO DEL ARTE

El reconocimiento de actividades es un campo que ha tomado mucha importancia en los últimos años, siendo este aplicado a diferentes áreas como la vigilancia de la seguridad pública, la realidad virtual y los juegos de ordenador, deportes, rehabilitación física, entre otros [30, 31]. A continuación se relacionan las diferentes metodologías que se han propuesto para el reconocimiento de actividades utilizando diferentes sistemas de captura de movimiento en ambientes no estructurados y empleando varias técnicas de fusión de datos y de aprendizaje de máquina para su clasificación.

En [9] se ha trabajado con los acelerómetros para la detección del movimiento. Se determinó la tasa de muestreo óptima de este tipo de señales mediante un análisis estadístico y de similitud en cinco conjuntos de datos en escenarios particulares manteniendo la exactitud de los sistemas de reconocimiento con la reducción de la frecuencia de muestreo. Por otro lado, en [23] se implementa K-means y las SVM para construir un libro de códigos con las posturas claves de las actividades ejecutadas con el Kinect®. Posteriormente se usan los HMM para identificar las diferentes combinaciones de posturas y así reconocer la actividad realizada. Las bases de datos con las que se implementó esta metodología constan de 18 y 12 actividades con eficiencias del 77.3 % y 77.6 % respectivamente utilizando 11 de los puntos articulados que entrega el Kinect®. Sin embargo, el Kinect® presenta problemas de oclusión en la realización de ciertas actividades obteniendo una clasificación errónea. Por lo tanto, es importante incluir otro tipo de sensores para el registro del movimiento permitiendo que la fusión de estos mejore la tasa de desempeño del clasificador. Además, el enfoque de detección de posturas solo puede ser aplicado en metodologías de estimación de puntos articulados lo que limita el uso a otras modalidades de sensores.

La fusión de múltiples sensores y de diferentes tipos es una tarea fundamental pero no trivial que afecta directamente el rendimiento de la aplicación, en este caso, la precisión en el reconocimiento de las actividades físicas. La fusión se puede realizar a nivel de datos, de características y de decisión [10]. En [27] se realiza la fusión de los datos provistos por los sensores IMU y Kinect® con el filtro de Kalman extendido (EKF) para la medición de los

movimientos del cuerpo. Además, se utiliza un filtro de Kalman previamente para corregir los problemas de oclusión por parte del Kinect®. También, se estima la varianza del ruido de los sensores relacionados con los movimientos de los brazos y piernas obteniendo un error cuadrático medio del $6,08^\circ$ con el método de fusión, el cual es menor en comparación con un sólo sensor y con el método de fusión sin el filtrado de los datos. Los sensores usados son el giroscopio, acelerómetro y el Kinect®. Por otro lado, en [1] se crea una base de datos multimodal de 5 actividades físicas con el registro sincronizado del Kinect®, IMU y EMG en una ventana de observación de tres segundos. En la clasificación de los movimientos primitivos se usan las SVM para cada sensor. En la identificación de las actividades físicas utiliza los HMM fusionando las etiquetas generadas por las SVM. El mejor desempeño se obtuvo con la configuración Kinect+ IMU del 98.81 % con la mínima cantidad de sensores. Sin embargo, en este trabajo no se experimento con diferentes tamaños de ventanas para la segmentación del movimiento, ya que esto influye directamente en el desempeño del clasificador. En [14] con la misma base de datos de [13] realiza la fusión a nivel de datos y de características utilizando técnicas de aprendizaje no supervisado como Spectral Clustering y K-means Clustering determinando la cantidad necesaria de movimientos primitivos. Para el reconocimiento de las actividades se usa las HMM con porcentaje de acierto del 70.21 % para todas las combinaciones de sensores, además, se realiza la selección de características relevantes utilizando PCA y Relief, donde Relief logra aportar los mejor resultados para el sistema de reconocimiento.

La segmentación del movimiento por medio de ventanas es una de las técnicas usadas para la identificación de los movimientos primitivos. De esta manera se realiza extracción de las características más relevantes de las ventanas para cada actividad. En [28] se propone un método estadístico para la representación de los movimientos primitivos y el reconocimiento de las actividades humanas utilizando BoF y un sistema multimodal conformado por una IMU de 6 grados de libertad. La IMU se compone de un acelerómetro triaxial, un giroscopio triaxial y un magnetómetro triaxial para el registro de las 9 actividades. En la construcción del vocabulario se utilizó un algoritmo de clustering no supervisado. En este sentido, cada cluster se consideró como un único movimiento primitivo construyendo de esta forma un

histograma con la distribución de los sub-movimientos mapeados en un vector de características globales. El modelo de aprendizaje para la identificación de las actividades fue las SVM con un kernel lineal y gaussiano con una eficiencia del 92.7%. En [29] se propone un método no supervisado de segmentación temporal basado en ganancias para encontrar los tiempos de transición entre las actividades humanas y las rutinas diarias con diferentes modalidades de sensores. El método propuesto se define por una función de costo donde se implementan dos métodos de optimización (TopDown y programación dinámica) para encontrar la mejor segmentación reduciendo la entropía. De este modo se usan cuatro bases de datos (una sintética y tres con datos reales) obteniendo un porcentaje de acierto del 71 %.

3. MARCO CONCEPTUAL

3.1. Unidad de Masa Inercial (IMU)

Es un sistema que realiza la medición de la aceleración lineal, velocidad angular y orientación utilizando acelerómetros, giroscopios y magnetómetros respectivamente. Este sensor realiza la medida integrada de los parámetros de movimiento en seis grados de libertad, determinando la posición de un objeto en movimiento dentro del espacio mediante las coordenadas (X, Y, Z) y los ángulos de rotación (pitch, yaw y roll), [21, 14].

3.2. Electromiografía (EMG)

Es una técnica para medir la actividad eléctrica que se produce por los músculos esqueléticos en la realización de un movimiento intencional (contracción o relajación de un conjunto de músculos medidos). Este método usa electrodos de superficie que miden la actividad muscular esquelética transmitida a la superficie de la piel de forma no invasiva [1].

3.3. Kinect

Dispositivo diseñado por la empresa Microsoft para la detección de movimiento y captar audio permitiendo la interacción con el usuario a través de comandos de gesto y voz. Este dispositivo integra varios componentes como sensor de profundidad, cámara RGB, un arreglo de micrófonos, información de los puntos articulados en 3D y una cámara IR, los cuales permiten obtener información de una escena [32].

3.4. Movimientos primitivos

Una actividad se puede representar como una secuencia continua de movimientos, esta secuencia de movimientos se segmenta en componentes más pequeñas denominadas movimientos primitivos. Lo cual permite facilitar la identificación del movimiento con algún modelo de aprendizaje. Los movimientos primitivos son utilizados en muchas aplicaciones, entre ellas se encuentra el reconocimiento automático de actividades físicas [21].

3.5. Fusión sensorial

La fusión de datos procedentes de múltiples fuentes sensoriales representa una solución efectiva para inferir información de alta calidad en señales con un alto contenido de ruido, en muestras aleatorias y sistemáticas del sensor. Estas señales son afectadas por el error, pérdida o inconsistencia de los datos.

La fusión sensorial se puede agrupar en:

- Fusión Competitiva

Implica la utilización de múltiples fuentes sensoriales equivalentes de información. Es utilizada para obtener redundancia y auto-calibración.

- Fusión Complementaria

Cada sensor captura diferentes aspectos de los fenómenos controlados. Se usa para mejorar la precisión y fiabilidad del sistema.

- Fusión Cooperativa

Cuando se necesitan las señales de múltiples sensores para obtener la información que no se puede conseguir con una solo tipo de sensor. [10].

3.6. Ventana

Consiste en seleccionar un pequeño subconjunto de datos de un conjunto más grande para su procesamiento y análisis. Los dos parámetros principales para la ventana son el tamaño y la cantidad de traslapes entre ventanas adyacentes.

Los algoritmos de ventana fija utilizan ventanas deslizantes de longitud fija, las cuales son sensibles a la longitud de esta y a la cantidad de traslapes subsecuentes. Generalmente el rango del traslape entre la ventana actual y la siguiente es del 0 % al 50 %. Los algoritmos de ventana de longitud variable emplean segmentos de datos que cambian de tamaño dinámicamente para ajustarse a los datos entrantes. Esta técnica tiene un alto costo computacional para determinar el tamaño de la ventana. Sin embargo, el tamaño de la ventana está más dirigido al movimiento subyacente, lo que potencialmente mejora la calidad de la segmentación [21].

3.7. Análisis de componentes principales (PCA)

Es una técnica no paramétrica que permite realizar la reducción de dimensionalidad de un conjunto de datos de alta dimensionalidad en un subespacio de menor dimensión, conservando las características importantes mediante la transformación del conjunto original a un nuevo conjunto de variables que están no correlacionadas. A este tipo de variables se les conoce como las componentes principales [22].

3.8. Extracción de características

Consiste en extraer información valiosa de los datos adquiridos de la señal de un sensor para la identificación de una entidad en particular. Algunas de estas señales son los fenómenos físicos, comportamientos, movimientos, entre otros donde se define un tamaño de ventana para la extracción de las características [1]. La extracción de características se clasifica en tres categorías (dominio en el tiempo, dominio en la frecuencia y dominio discreto) donde se genera un vector de características utilizadas por el modelo de clasificación en la etapa de entrenamiento [25].

3.9. Clasificación

Hace parte del área de aprendizaje de máquina. Consiste en estimar un modelo matemático que separe o identifique una entidad utilizando la información suministrada. La clasificación se puede realizar de forma supervisada, no supervisada, semi-supervisada, por aprendizaje por refuerzo, transducción o aprendizaje multi-tarea, en línea, activo. Algunas técnicas de aprendizaje utilizadas para el proceso de clasificación son las máquinas de vectores de soporte (SVM), los modelos ocultos de Markov (HMM), las redes neuronales (RNA), los vecinos más cercanos (KNN), entre otros.

3.9.1. Clasificación Supervisada

Son técnicas de aprendizaje, donde se recibe un conjunto de ejemplos etiquetados como datos de entrenamiento y se hacen predicciones para todos los puntos. Estas técnicas en general

buscan la estimación de los parámetros de un hiperplano de separación que permitan establecer un límite entre las etiquetas a asignar [33].

3.9.2. Clasificación No Supervisada

Son metodologías de agrupamiento donde el aprendizaje se realiza exclusivamente con los datos de entrenamiento no etiquetados con el fin de hacer las predicciones para todos los puntos no vistos. Se distingue de la clasificación supervisada, por el hecho de que no hay un conocimiento a priori del comportamiento de las observaciones. La finalidad es agrupar entidades que pertenezcan a un mismo conjunto [33].

3.10. Análisis de Desempeño de un clasificador

Se considera un problema de clasificación binaria donde se tienen indicadores que permiten realizar la validación de un conjunto de datos con los que se prueba el algoritmo. Estos indicadores se definen como [34]:

- Verdaderos Positivos (TP): Son los ejemplos que pertenecen a la clase y el clasificador los estimó que pertenecen a la clase.
- Verdadero Negativos (TN): Son los ejemplos que no pertenecen a la clase y el clasificador los estimó que no pertenecen a la clase.
- Falsos Positivos (FP): Son los ejemplos que no pertenecen a la clase y el clasificador los estimó que pertenecen a la clase.
- Falsos Negativos (FN): Son los ejemplos que pertenecen a la clase y el clasificador los estimó que no pertenecen a la clase.

Con estos indicadores se pueden obtener una serie de métricas para estimar el rendimiento del clasificador como [34]:

- Sensibilidad o recall: Es la proporción de verdaderos positivos de un clasificador y esta definida como:

$$Sensibilidad = \frac{TP}{TP + FN} \quad (1)$$

- Proporción de falsa alarma (FA): Es la proporción de falsos positivos de un clasificador y esta definida como:

$$FA = \frac{FP}{FP + TN} \quad (2)$$

- Especificidad: Que tan efectivo es el clasificador identificando las etiquetas negativas y esta definida como:

$$Especificidad = \frac{TN}{FP + TN} \quad (3)$$

4. MARCO TEÓRICO

4.1. Máquinas de vectores de soporte (SVM) [1, 2]

Las máquinas de vectores de soporte surgen de los trabajos desarrollados por Vapnik sobre la teoría del aprendizaje estadístico para resolver problemas de clasificación binaria, regresión, agrupamiento y multclasificación. Las SVMs pertenecen a la categoría de los clasificadores lineales donde el objetivo es realizar la separación de los datos por medio de hiperplanos que dividan y maximicen el margen entre las clases en este espacio.

Dado un conjunto de datos de dimensionalidad D :

$$\{\mathbf{x}_i, y_i\} \text{ donde } i = 1 \dots m, y_i \in \{-1, 1\}, \mathbf{x} \in \mathbb{R}^D$$

La ecuación de un hiperplano viene dada por:

$$h = \mathbf{w}^T \mathbf{x} + b \quad (4)$$

Donde:

- \mathbf{w} es un vector de pesos con los coeficientes en el hiperplano.
- \mathbf{x} es un vector de datos.
- $\frac{b}{\|\mathbf{w}\|}$ es la distancia perpendicular desde el hiperplano hasta el punto de origen.
- y_i son la etiquetas de los datos.

Los vectores de soporte son los ejemplos más cercanos al hiperplano de separación y el objetivo de las SVM es orientar este hiperplano de tal manera que estén lo más lejos posible de los miembros más cercanos de ambas clases. Por consiguiente, se busca estimar una función tal que para una entrada en \mathbb{R}^D produzca una salida en $\{\pm 1\}$ donde:

$$f : \mathbb{R}^D \rightarrow \{+1, -1\}$$

La distancia entre el hiperplano de separación y los puntos más cercanos es la mitad de la distancia entre estos hiperplanos de soporte, que es la misma que la diferencia entre las

distancias al origen del punto más cercano en cada uno de los hiperplanos de soporte, esto se conoce como el margen de las SVM. Sin embargo, los datos no necesariamente son linealmente separables por lo que las SMV abordan el problema de clasificación relajando el concepto de margen introduciendo un sesgo ξ_i , donde la función de costo o función fitness se define como:

$$c(y, f(x, \mathbf{w})) = \begin{cases} 0 & \text{si } y = \text{sgn}(w^T \phi(x) + b) \\ 1 & \text{si } y \neq \text{sgn}(w^T \phi(x) + b) \end{cases} \quad (5)$$

Donde, $\phi(x) : \mathbb{R} \rightarrow H$ es una transformación del conjunto de entrada a un conjunto con igual o mayor dimensión. Esto supone que facilita la separación de las clases. El vector \mathbf{w} define el hiperplano de separación en el espacio $\phi(x)$ y b representa el bias o sesgo con respecto al origen de coordenadas. En este sentido, con la relajación del margen de las SVM se busca que los puntos de datos que están en el lado incorrecto del límite del margen tengan una penalización que aumente con la distancia desde el mismo. Por lo tanto, se trata de reducir el número de clasificaciones erróneas adaptando la función de costo a la solución de un problema de optimización cuadrática (QP-Quadratic Problem) que viene dado por:

$$\min_{\mathbf{w}, b, \xi_i} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^m \xi_i \quad (6)$$

$$\text{sujeto a } y_i (\mathbf{w}^T \phi(x_i) + b) \geq 1 - \xi_i \quad (7)$$

$$\xi_i \geq 0, \forall_i = 1, \dots, m \quad (8)$$

Donde:

- C controla la compensación entre la penalización de holgura variable y el tamaño del margen, es decir es una constante de regularización.
- x_i es una muestra del conjunto de entrenamiento con su respectiva etiqueta.
- ξ_i variable slack que permite relajar el margen de clasificación de la SVM.

Reformulando como un lagrangiano, minimizando con respecto \mathbf{w}, b , y ξ_i y maximizando con respecto a $\boldsymbol{\alpha}$, se tiene:

$$\max_{\alpha_i} L_D = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j \phi^T(x_i) \phi(x_j) \quad (9)$$

$$\text{sujeto a } \sum_{i=1}^m \alpha_i y_i = 0 \quad (10)$$

$$0 \leq \alpha_i \leq C, \quad \forall_i = 1, \dots, m \quad (11)$$

4.1.1. Funciones Kernel [3]

Hasta el momento se ha asumido que los datos son separable o cuasi-separables, por lo que los hiperplanos se han definido como funciones lineales en el espacio \mathbf{x} de los datos. Sin embargo, no en todos los casos es convenientes que las funciones de los hiperplanos sean definidos con funciones lineales, ya que se pueden tener datos no linealmente separables donde es conveniente utilizar una función Kernel ($K(\mathbf{x}, \mathbf{x}')$). Una función $K : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}$ que se asigna a cada elemento del espacio de entrada \mathbb{X} un valor real correspondiente al producto escalar de las imágenes de dichos elementos en un nuevo espacio \mathcal{F} (espacio de características), es decir:

$$K(\mathbf{x}, \mathbf{x}') = \langle \Phi(\mathbf{x}) \Phi(\mathbf{x}') \rangle = (\phi_1(\mathbf{x}) \phi_1(\mathbf{x}') + \dots + \phi_m(\mathbf{x}) \phi_m(\mathbf{x}')) \quad \forall \mathbf{x}, \mathbf{x}' \in \mathbb{X} \quad (12)$$

Donde $\Phi : \mathbb{X} \rightarrow \mathcal{F}$

Por lo tanto la función Kernel es muy útil para problemas de clasificación/regresión que no son linealmente separables/regresivos en el espacio de las \mathbf{x} entradas, la cuales podrían estar en un espacio de características de alta dimensionalidad dado un mapeo adecuado $\mathbf{x} \rightarrow \phi(\mathbf{x})$.

A continuación se muestran algunas funciones Kernel:

- Kernel lineal:

$$K(\mathbf{x}, \mathbf{x}') = \langle \mathbf{x}, \mathbf{x}' \rangle \quad (13)$$

- Kernel polinómico de grado p :

$$K_p(\mathbf{x}, \mathbf{x}') = [\gamma \langle \mathbf{x}, \mathbf{x}' \rangle + \tau]^p \quad (14)$$

- Kernel gaussiano:

$$K(\mathbf{x}, \mathbf{x}') = \exp\left(-\gamma \|\mathbf{x} - \mathbf{x}'\|^2\right), \quad \gamma = \frac{1}{2\sigma^2} \quad (15)$$

4.1.2. SVM Multiclase [4]

Originalmente las SVM son utilizados para problemas de clasificación binario. Sin embargo, se suelen encontrar problemas donde se tienen más de dos clases, en este caso las SVM tienen varias metodologías:

- **Uno contra todos (one versus all):**

Dado un grupo de n datos de entrenamiento donde existen l clases ($l > 2$), se puede tener un grupo de m clasificadores binarios (donde $m = l$), cada uno entrenado para separar una clase del resto de las clases existentes ($l - 1$), para lo cual se construye una matriz de descomposición D de la siguiente forma:

$$D_{i,j} = \begin{cases} +1 & \text{si } n \in n_j \\ -1 & \text{si } n \in n_i \end{cases} \quad (16)$$

Donde n_j son los datos pertenecientes a la clase $j \in \{1, 2, \dots, l\}$ que se etiquetan con $t_j = +1$ y el resto de datos $n_i = n - n_j$ se etiquetan con $t_r = -1$.

- **Uno contra uno (one versus one):**

Consiste en entrenar $l(l-1)/2$ clasificadores binarios, donde l es el número total de clases. El entrenamiento de cada clasificador binario se realiza con dos de las clases del grupo de datos de entrenamiento, para lo cual se construye una matriz de descomposición D de la siguiente forma:

$$D_{i,j} = \begin{cases} +1 & \text{si } n \in n_j \\ -1 & \text{si } n \in n_i \\ 0 & \text{si } n \in n_r \end{cases} \quad (17)$$

Donde n_j son los datos pertenecientes a la clase $j \in \{1, 2, \dots, l\}$ que se etiquetan con $t_j = +1$, n_i son los datos pertenecientes a la clase $i \in \{1, 2, \dots, l\}$ donde $i \neq j$ se etiquetan con $t_i = -1$, y el resto de datos $n_r = n - n_j - n_i$ no participan en el entrenamiento donde la etiqueta es $t_r = 0$.

4.1.3. Método 1 de penalización: C-Clasificación de los Vectores de Soporte(C-SVC) [5]

Método de penalización resolviendo para un $C > 0$

$$\min \tau(\mathbf{w}, \boldsymbol{\xi}) = \frac{1}{2} \|\mathbf{w}\|^2 + \frac{C}{m} \sum_{i=1}^m \xi_i \quad (18)$$

Sujeto a las restricciones 7 y 8 para una constante positiva C adecuada, este enfoque minimiza aproximadamente el lado derecho del límite. Sin embargo, si muchos de los valores de ξ_i son altos debido a que las clases que se van a separar se superponen fuertemente, entonces $\sum_{i=1}^m \xi_i$ puede ser significativamente mayor que la fracción de errores del margen. Por lo tanto, no hay garantía de que el hiperplano se generalizará bien.

Para el caso separable,

$$\mathbf{w} = \sum_{i=1}^m \alpha_i y_i x_i \quad (19)$$

Donde los coeficientes de α_i distintos de cero solo pueden ocurrir si el ejemplo (x_i, y_i) cumple con la restricción 7. Los coeficientes α_i se calculan resolviendo un problema de optimización cuadrática (QP-Quadratic Problem) que viene dado por:

$$\max_{\alpha_i} L_D = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j \phi^T(x_i) \phi(x_j) \quad (20)$$

$$\text{sujeto a } \sum_{i=1}^m \alpha_i y_i = 0$$

$$0 \leq \alpha_i \leq \frac{C}{m}, \forall i = 1, \dots, m$$

4.1.4. Método 2 de penalización: SVM binaria ponderada [6]

Método de penalización considerando un problema de clasificación con un conjunto de clases desbalanceadas, siendo D un conjunto de datos de $\{f_i, y_i\}_{i=1}^N$, donde:

- $f_i \in R^d$ es la i -ésima muestra y $y_i \in \{\omega_1, \dots, \omega_c\}$ son la etiquetas pertenecientes a cada clase.
- N es el número total de datos de entrenamiento.

Sea N_j el número de datos de entrenamiento que pertenecen a la clase ω_j y $D_{ij} := \{f_k, y_k\}_{f_k \in \omega_i \cup \omega_j}$ el subconjunto de D formado por los datos pertenecientes a las clases ω_i y ω_j . Las SVM binaria ponderada realiza la clasificación de las clases (ω_i y ω_j) del conjunto de datos de entrenamiento D_{ij} resolviendo el siguiente problema de optimización:

$$\max_{w^{ij}, b^{ij}, \xi^{ij}} \frac{1}{2} (w^{ij})^T w^{ij} + C_i^{ij} \sum_{f_k \in \omega_i} \xi_k^{ij} + C_j^{ij} \sum_{f_k \in \omega_j} \xi_k^{ij} \quad (21)$$

$$\text{sujeto a } w^{ij} \cdot \phi^{ij}(f_k) + b^{ij} \geq 1 - \xi_k^{ij}, \text{ si } f_k \in \omega_i$$

$$w^{ij} \cdot \phi^{ij}(f_k) + b^{ij} \leq -1 - \xi_k^{ij}, \text{ si } f_k \in \omega_j$$

$$\xi_k^{ij} \geq 0 \forall k$$

Donde:

- ϕ^{ij} es una función no lineal que mapea los vectores de características en un espacio euclidiano de alta dimensionalidad H .
- $w^{ij} \in H$ y $b^{ij} \in R$ son los parámetros que determinan la separación óptima del hiperplano.
- $\xi_k^{ij} \in R$ es la variable slack.
- Para este caso, se tiene diferentes parámetros de regularización en 21 (C_i^{ij} y C_j^{ij}) para las clases ω_i y ω_j respectivamente:

$$w^{ij} \cdot \phi^{ij}(f) + b^{ij} = 0 \quad (22)$$

Con la ecuación 23 se impone una mayor penalización en los errores de clasificación que se tienen en las muestras de la clase de menor tamaño debido al desbalanceo que existen entre las clases.

$$C_i^{ij}/C_j^{ij} = N_i/N_j \quad (23)$$

La siguiente formulación dual permite encontrar la solución numérica de 21

$$\max_{\alpha_i} L_D = \sum_{f_k \in D_{ij}} \alpha_k^{ij} - \sum_{f_k \in D_{ij}} \sum_{f_r \in D_{ij}} \alpha_k^{ij} \alpha_r^{ij} y_k y_r K^{ij}(f_k, f_r) \quad (24)$$

$$\text{suje}to a \sum_{f_k \in D_{ij}} y_k \alpha_k^{ij} = 0$$

$$0 \leq \alpha_k^{ij} \leq C_i^{ij}, \text{ si } f_k \in \omega_i, 0 \leq \alpha_k^{ij} \leq C_j^{ij}$$

Donde α_k^{ij} son los multiplicadores de Langrange para la k -ésima muestra y $K^{ij}(f_k, f_r) = \phi^{ij}(f_k) \cdot \phi^{ij}(f_r)$ el la función kernel.

4.2. Modelos Ocultos de Markov (HMM) [7]

Los modelos ocultos de Markov son una herramienta estadística empleada en el moldeamiento de series de tiempo. Por lo tanto, es un método doblemente estocástico el cual tiene un proceso oculto (secuencia de estados finitos) que sólo puede ser asociado probabilísticamente con otro proceso estocástico observable, produciendo la secuencia de características que se pueden observar. El principal objetivo de los HMM es encontrar el modelo que mejor explica una secuencia de observaciones dadas dentro de un conjunto de modelos.

Los HMM está definido por:

- Los estados ocultos $h_i \in H$ con $i = 1, 2, \dots, m$, donde m define el tamaño del vector de estados ocultos H_t .
- El vector de las posibles observaciones $X = \{x_1, x_2, \dots, x_n\}$, donde n define la longitud del vector de observaciones teniendo en cuenta $m \leq n$.
- La distribución de los estados iniciales $\pi = (\pi_1, \pi_2, \dots, \pi_m)^T$, donde el primer estado inicial h_0 es la probabilidad marginal (variable independiente e idénticamente distribuida).

Por lo tanto:

$$[\pi]_i = p(h_i) \quad (25)$$

Para hacer la transición entre los estados ocultos h_t de los HMM se necesita una matriz de probabilidad de transición de estado $T \in \mathbb{R}^{m \times m}$. Definida como:

$$[T]_{j,i} = p(h_j|h_i) \quad (26)$$

Además, se tiene la matriz de probabilidad de salida O que representa la distribución de probabilidad para muestrear un símbolo observable desde un estado oculto. Definida como:

$$[O]_{k,j} = p(x_k|h_j) \quad (27)$$

Teniendo en cuenta la ecuaciones 25, 26 y 27 se construye λ , que es un arreglo donde el estado oculto del instante t depende únicamente del último estado oculto. Por lo tanto, los HMM se define como un 3-tupla $\lambda = (T, O, \pi)$. Donde:

$$\forall t : p(h_t|h_{1:t-1}) = p(h_t|h_{t-1}) \quad (28)$$

$$\forall t : p(x_t|h_t) \quad (29)$$

En el que $o_j(x_k)$ es la probabilidad de emitir un símbolo de observación en el estado h_j .

En los HMM se presentan tres tipos de problemas fundamentales:

1. Problema de Evaluación: El cual consiste en determinar la probabilidad $p(X|\lambda)$ de que el modelo genere las observaciones, dado un modelo λ y una secuencia de observaciones X .
2. Problema de Decodificación: Donde la idea es determinar la secuencia de estados ocultos H más probable en el modelo, que se generó a partir de una secuencia de observación X .
3. Problema de Aprendizaje: En el cual se debe encontrar un algoritmo que ajuste los parámetros del modelo λ para maximizar la probabilidad conjunta $p(X|\lambda)$.

5. CAPTURA DE DATOS DEL SISTEMA MULTIMODAL

En diversos trabajos el análisis de actividades físicas humanas constituye un tarea fundamental aplicadas a diferentes áreas. Por consiguiente, es importante utilizar un sistema de medición que permita la captura del movimiento. En este sentido, se hace necesario tener una base de datos anotada que permita evaluar en igualdad de condiciones los comportamientos estáticos y dinámicos de las señales registradas por las diferentes modalidades de sensores. Actualmente existen varios tipos de sensores que permiten el registro del movimiento como es el Kinect®, MoCap (Motion Capture System), IMU y EMG de los cuales se tienen bases de datos anotadas estandarizadas usadas para el reconocimiento de actividades físicas humanas, como lo son:

- **KARD (Kinect Activity Recognition Dataset) [35]**

Contiene el registro de 18 actividades utilizando el Kinect® como dispositivo de captura. Los datos adquiridos de este dispositivo son las imágenes de profundidad, los vídeos en RGB y los puntos articulados del esqueleto. Cada acción fue realizada por 10 sujetos diferentes con una repetición de 3 veces y capturados a una resolución de 640×480 píxeles a $30fps$.

- **OPPORTUNITY Activity Recognition Data Set [36]**

Es una amplia base de datos que contiene el registro de 12 personas con más de 27000 actividades de la vida cotidiana ejecutadas en un ambiente con 72 sensores de diferentes modalidades. Los sensores utilizados son de aceleración, localización, ambientales, IMU's y de velocidad.

- **REALDISP (REAListic sensor DISPlacement) [37]**

Esta base de datos tiene el registro de 33 actividades enfocadas al monitoreo del ejercicio físico con 17 personas usando 9 sensores IMU y estableciendo tres escenarios donde se modifica la ubicación de estos.

- **NTU RGB+D (Nanyang Technological University's Red Blue Green and**

Depth information) [38]

Tiene más de 56000 vídeos de 40 personas con 60 clases de acciones usando 3 sensores Kinect One®. Además, cuenta con el registro sincronizado de los vídeos en RGB a una resolución de 1920×1080 , las secuencias de los mapas de profundidad, los vídeos en infrarrojo de 512×480 y los datos del esqueleto en 3D con las 25 articulaciones principales del cuerpo.

■ **Berkeley MHAD (Berkeley Multimodal Human Action Database)[39]**

Es una base de datos multimodal que tiene el registro de 11 acciones ejecutadas por 12 personas con 5 repeticiones. Para este caso, se utilizaron cinco sistemas diferentes de captura de movimiento (MoCap, 2 Kinect One®, 2 clusters con cámaras de visión estereoscópica, dos clusters con 4 cámaras para captura multi-vista y 6 acelerómetros inalámbricos) produciendo un vídeo de 82 minutos con 660 secuencias del tiempo total de la grabación para cada actividad.

En otras publicaciones enfocadas al reconocimiento de actividades físicas han desarrollado diferentes bases de datos con algunos de los sensores mencionados con anterioridad. Sin embargo, no siempre están públicas o son de libre descarga como se menciona en [31, 40, 18, 41]. Además, la captura del movimiento se realiza generalmente con el dispositivo Kinect® y las IMU sin incluir los sensores electromiográficos. Por otro lado, en [1] se construyó una base de datos anotada multimodal con el registro sincronizado de tres modalidades de sensores (Kinect®, IMU y EMG). Esta base de datos consta de 5 actividades ejecutadas por 16 personas en 3 repeticiones y se encuentra disponible en <https://goo.gl/6F82wd>. No obstante, se tienen pérdidas de información durante la captura de datos de los IMUs y EMGs. La lista de actividades es reducida con movimientos primitivos totalmente separables entre sí. Por lo tanto, no se puede hacer un análisis del desempeño del método con actividades que tengan características similares como caminar, trotar y correr. Como consecuencia, es necesario crear una base de datos de 10 actividades con algunas retos en la clasificación de actividades como caminar y trotar, sentarse y levantarse, entre otras. Para la adquisición de datos se utilizó

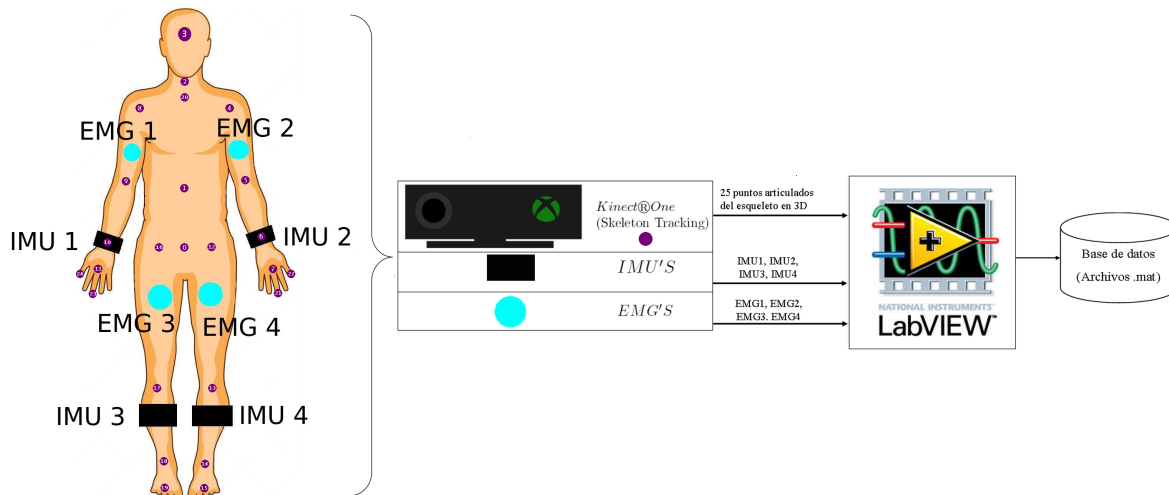
el software de LabView que permite el registro sincronizado de las señales provistas por los sensores como lo señala [13], la cual consta de cuatro etapas:

1. Captura de los puntos articulados del cuerpo utilizando el Skeleton Tracking provisto por el Kinect One®.
2. Captura sincronizada de la red de sensores IMU (4 sensores ubicados en el miembro superior e inferior del cuerpo).
3. Captura sincronizada de la red de sensores EMG (4 sensores ubicados en el miembro superior e inferior del cuerpo).
4. Almacenamiento del registro sincronizado de los sensores en una estructura binaria .mat utilizado por entornos como Matlab y Octave.

5.1. Base de datos anotada de actividades físicas humanas

En la figura 1 se visualiza el sistema de captura de movimiento propuesto por *Andrés et al* [1] indicando la ubicación de cada uno de los sensores en las persona.

Figura 1. Esquema de la distribución de los sensores y del sistema de captura sincronizado.



A continuación se listan las especificaciones de cada uno sensores utilizados:

- Para el dispositivo Kinect One®

Entrega 25 puntos de las articulaciones principales del cuerpo humano mediante la cámara de profundidad que tiene este dispositivo a una frecuencia de 15 muestras por segundo. Teniendo en cuenta que en [1] se utilizó la primera versión del Kinect® con el registro de sólo 14 puntos de las articulaciones a una frecuencia de 30 muestras por segundo, con un menor rango de cobertura en comparación a la segunda versión de este dispositivo [42, 43, 44, 45].

- **Para la red de sensores IMU**

Se utilizan cuatro acelerómetros tri-axiales ADLX330 de 3 grados de libertad (ax, ay, az) a una frecuencia de muestreo de 30 muestras por segundo. Además, se calcularon los ángulos de rotación *Roll* y *Pitch* para obtener las medidas de un sensor IMU.

- **Para la red de sensores EMG**

Consta de ocho electrodos de superficie en conexión diferencial a una frecuencia de muestreo de 2 [KHz] .

El diseño de los diagramas de bloques para la adquisición de datos de las redes de sensores IMU y EMG se realizó en LabView utilizando los VIs implementados en [1]. Sin embargo, para el Kinect One® fue necesario desarrollar los VIs para el registro de los 25 puntos articulados que entrega este dispositivo. En las figuras 2, 3 y 4 se presenta el diseño de los algoritmos para la adquisición de los datos entregados por las redes de sensores y del dispositivo Kinect One®, y en la figura 5 se observa la interfaz gráfica de captura para cada actividad.

Figura 2. Adquisición del dispositivo Kinect One®.

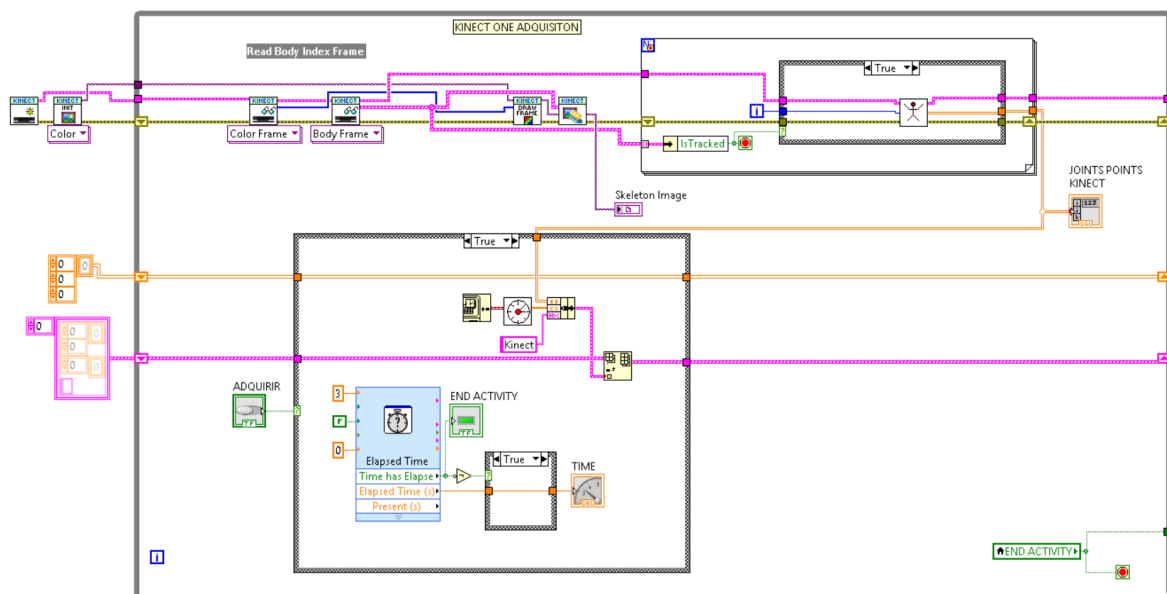


Figura 3. Adquisición de la red de sensores EMG.

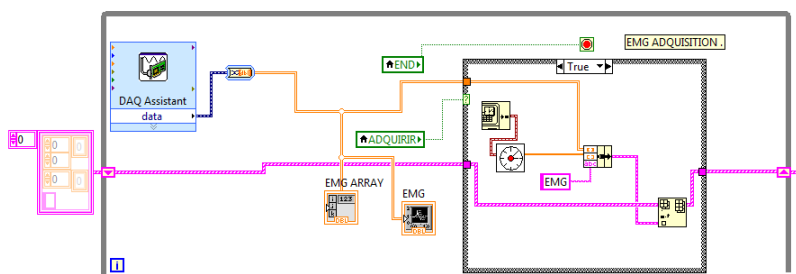


Figura 4. Adquisición de la red de sensores IMU.

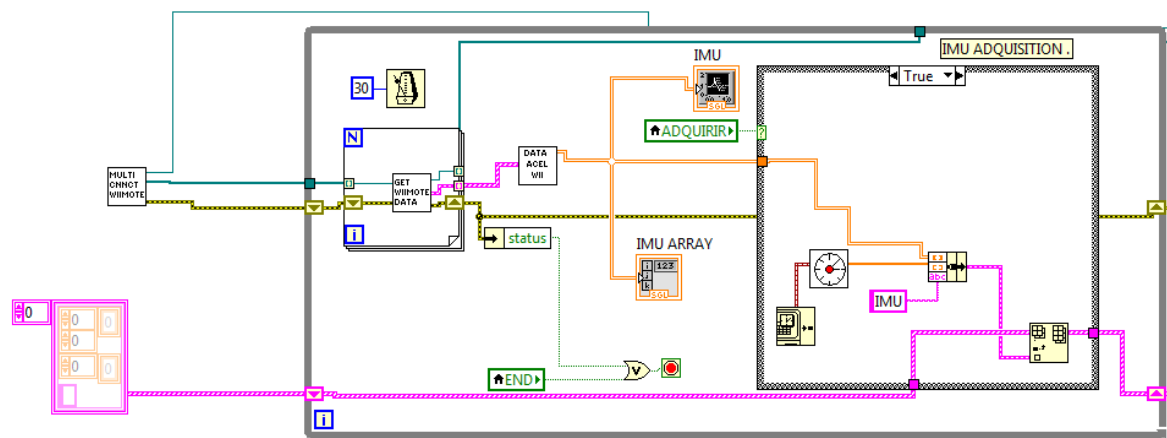
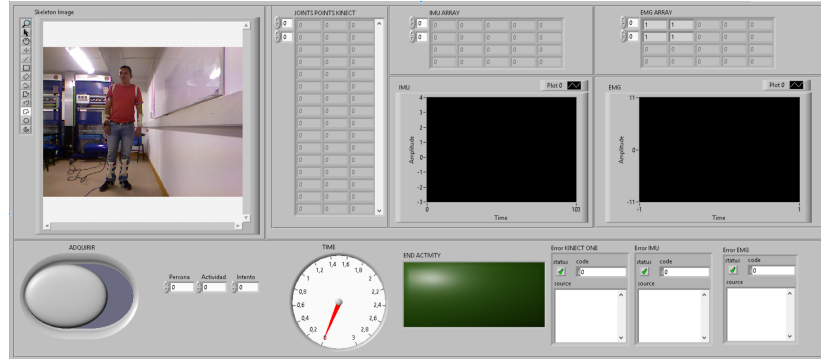


Figura 5. Diseño de la interfaz gráfica de captura de actividades.



En este caso, el almacenamiento sincronizado de los datos para las tres modalidades de sensores se realiza a través de una bandera booleana que activa la grabación simultánea (muti-hilo) de los datos con una ventana de observación de tres segundos por actividad. La codificación de cada acción se determinó con la nomenclatura Persona k _Actividad i _Intento j .mat. Las actividades que se establecieron para la creación de la base de datos hacen parte de la vida cotidiana de las personas y que se han trabajado en [8, 38, 21, 39, 13, 46] para el reconocimiento de actividades físicas humanas. En la tabla 1 se listan las 10 actividades ejecutadas por 12 personas (8 hombres y 4 mujeres con 5 repeticiones) con su respectiva etiqueta.

Tabla 1. Lista de actividades.

| Actividad | Etiqueta |
|-------------------------|----------|
| Quedarse quieto | 1 |
| Agacharse y levantarse | 2 |
| Saltar | 3 |
| Levantar mando derecha | 4 |
| Trotar | 5 |
| Levantar mano izquierda | 6 |
| Levantar las dos manos | 7 |
| Caminar | 8 |
| Sentarse | 9 |
| Levantarse | 10 |

En la figura 6 se observa la distribución de cada uno de los archivos .mat de la base de datos. Cada archivo contiene una celda de 1×3 con el registro de la información de los tres segundos. Cada celda almacena otra celda de 1×3 correspondiente a los datos registrados por

cada sensor.

Figura 6. Estructura de la base de datos.

| Base de Datos | | | | | | | | | |
|-----------------------------------|-----------------------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|------------------------------------|
| Actividad 1 | Actividad 2 | Actividad 3 | Actividad 4 | Actividad 5 | Actividad 6 | Actividad 7 | Actividad 8 | Actividad 9 | Actividad 10 |
| Persona1_Actividad1_Intento1.mat | Persona1_Actividad2_Intento1.mat | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | Persona1_Actividad10_Intento1.mat |
| Persona1_Actividad1_Intento2.mat | Persona1_Actividad2_Intento2.mat | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | Persona1_Actividad10_Intento2.mat |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| Persona12_Actividad1_Intento5.mat | Persona12_Actividad2_Intento5.mat | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | Persona12_Actividad10_Intento5.mat |

5.2. Base de datos anotada de movimientos primitivos

Utilizando la base de datos de actividades físicas de la sección 5.1 se construyó esta base de movimientos primitivos segmentando la señal de cada actividad con la secuencia de sub-movimientos para cada una de las acciones. El conjunto de datos recolectados en 3 segundos se divide en N ventanas. El tamaño de la ventana se determinó en función de la frecuencia de muestreo de cada uno de los sensores. En la sección 7.1 se describen los tamaños de las ventanas de observación para cada sensor. Estas señales se almacenan en una archivo .mat codificado de la siguiente manera:

$$Base \{Ejemplo\} \{Segundo\} \{Sensor\} \{Segmento\}$$

Donde:

- *Ejemplo* corresponde a una celda de 1×600 , con la información de cada actividad de la base de datos.
- *Segundo* corresponde a una celda de 1×3 , con la información de las tres modalidades de sensores (Kinect One®, IMU y EMG) durante los 3 segundos.
- *Sensor* corresponde a una celda de 1×3 , con la información de cada modalidad de sensor: Kinect One®, IMU y EMG.
- *Segmento* son los N arreglos correspondientes a cada movimiento primitivo almacenado.

La etiquetación de la base de datos se realizó de forma manual. De esta manera se observó la distribución espacial de las diferentes posturas suministradas por el Kinect One® durante los

tres segundos de grabación estableciendo la separación entre cada uno de los sub-movimientos. Para la selección de cada movimiento primitivo se tuvo en cuenta que estos fueran representativos para las diferentes actividades. Además, se consideraron las fases presentes en el estudio del análisis cinemático del miembro superior e inferior del cuerpo humano usados en [47, 48, 49, 50]. En la tabla 2 se muestran los movimientos primitivos que se determinaron para cada una de las actividades de la tabla 1.

Tabla 2. Lista de movimientos primitivos.

| Actividad | Movimiento primitivos |
|-------------------------|--|
| Quedarse quieto | Reposo |
| Agacharse y levantarse | Reposo, Medio Agachado, Totalmente Agachado |
| Saltar | Reposo, Suspendido en el aire $1/4$ |
| Levantar mano derecha | Reposo, Mano derecha levantada a $1/4$, Mano derecha levantada a $3/4$ |
| Trotar | Reposo, Pierna izquierda hacia adelante con rodilla flexionada, Pierna derecha hacia adelante con rodilla flexionada |
| Levantar mano izquierda | Reposo, Mano izquierda levantada a $1/4$, Mano izquierda levantada a $3/4$ |
| Levantar las dos manos | Reposo, Manos levantadas a $1/4$, Manos levantadas a $3/4$ |
| Caminar | Reposo, Paso adelante pie izquierdo, Paso adelante pie derecho |
| Sentarse | Reposo, Medio sentado, Sentado |
| Levantarse | Reposo, Medio sentado, Sentado |

A continuación, se describe el procedimiento implementado para la asignación de etiquetas a cada uno de los movimientos primitivos según la actividad. También, se determinan los puntos articulados del esqueleto y los umbrales para cada caso:

■ Actividad 2

Se utiliza la coordenada y del punto articulado 3 mostrado en la figura 1 que corresponde a la cabeza la cual se normalizo con respecto al valor máximo. Se establecieron tres umbrales para cada uno de los movimientos primitivos mostrados en la figura 8.

■ Actividad 3

Se usa la metodología de la actividad 2 pero con dos umbrales como se observa en la figura 8.

■ Actividad 4

Con la coordenada y del punto articulado 10 que se observa en la figura 1 que es la muñeca de la mano derecha se normaliza con respecto al valor máximo. En este caso, se plantearon tres umbrales de separación como se evidencia en la figura 8.

■ Actividad 5

Se utilizaron los puntos articulados 12, 13, 14, 16, 17, 18 mostrados en la figura 1 que coinciden con la cadera, rodilla y tobillo de la pierna izquierda y derecha respectivamente. Para este caso, se calculan los ángulos del movimiento de flexión y extensión de la rodilla derecha e izquierda (θ_I y θ_D) como lo propone [51]. Los vectores se determinaron con los puntos articulados 12, 13 y 14 para la pierna izquierda ($\overrightarrow{V_1I}$ y $\overrightarrow{V_2I}$) y 16, 17, 18 para la pierna derecha ($\overrightarrow{V_1D}$ y $\overrightarrow{V_2D}$) mostrados en la figura 7. Para el calculo de los ángulos se implementaron las ecuaciones 30, 31, 32, 33 y 34 . En esta actividad se propusieron tres umbrales que se visualizan en la figura 8 con la variación de $\Delta\theta$ que es resultado de la diferencia entre los ángulos θ_D y θ_I (ecuación 35).

$$\overrightarrow{V_1I} = P_{12} - P_{13} = (x_{12}, y_{12}, z_{12}) - (x_{13}, y_{13}, z_{13}) \quad (30)$$

$$\overrightarrow{V_2I} = P_{14} - P_{13} = (x_{14}, y_{14}, z_{14}) - (x_{13}, y_{13}, z_{13}) \quad (31)$$

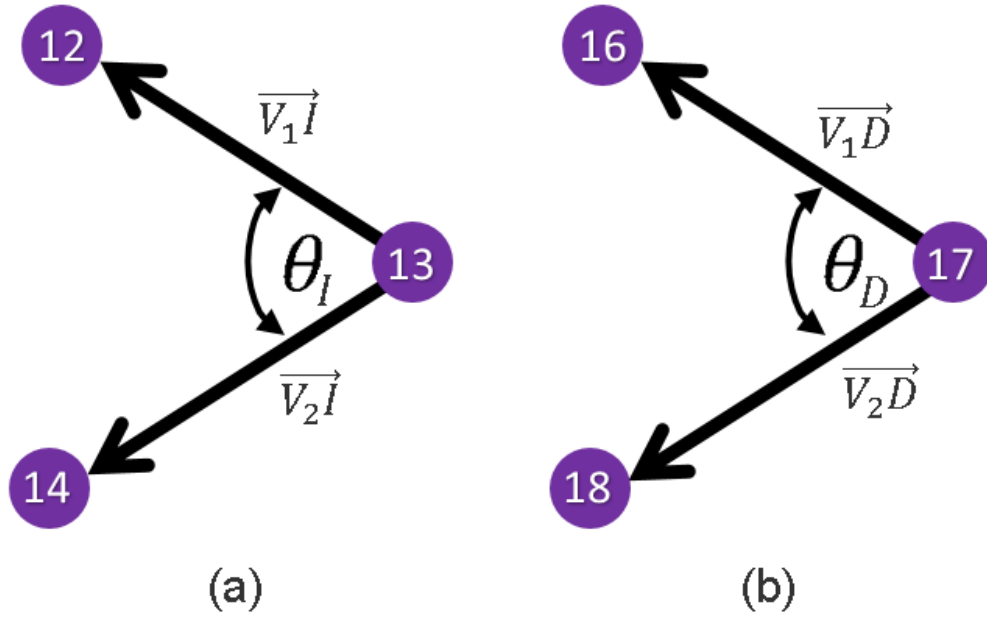
$$\overrightarrow{V_1D} = P_{16} - P_{17} = (x_{16}, y_{16}, z_{16}) - (x_{17}, y_{17}, z_{17}) \quad (32)$$

$$\overrightarrow{V_2D} = P_{18} - P_{17} = (x_{18}, y_{18}, z_{18}) - (x_{17}, y_{17}, z_{17}) \quad (33)$$

$$\theta_j = \arccos \left(\frac{V_{1j} \cdot V_{2j}}{|V_{1j}| |V_{2j}|} \right), \forall_j = D, I \quad (34)$$

$$\Delta\theta = \theta_D - \theta_I \quad (35)$$

Figura 7. Ángulo de flexión y extensión de la rodilla. a) Pierna izquierda, b) Pierna derecha.



■ Actividad 6

Se usa la metodología de la actividad 4 pero con el punto articulado 6 que pertenece a la muñeca de la mano izquierda, como se observa en la figura 8.

■ Actividad 7

Se utiliza la metodología de las actividades 4 y 6 estableciendo tres umbrales como se visualiza en la figura 8.

■ Actividad 8

Se usa la coordenada z de los puntos articulados del tobillo izquierdo y derecho 14 y 18

para realizar la diferencia con la coordenada z del centro de la cadera (punto articulado 0) como observan en las ecuaciones 36 y 37. Por consiguiente, se normalizan ambas medidas con respecto al mayor valor. Con la ecuación 38 se determinan los tres umbrales que se observan en la figura 8.

$$Z_D = Z_{\text{tobillo}_D} - Z_{\text{centro cadera}} \quad (36)$$

$$Z_I = Z_{\text{tobillo}_I} - Z_{\text{centro cadera}} \quad (37)$$

$$\Delta Z = Z_{n_D} - Z_{n_I} \quad (38)$$

Donde:

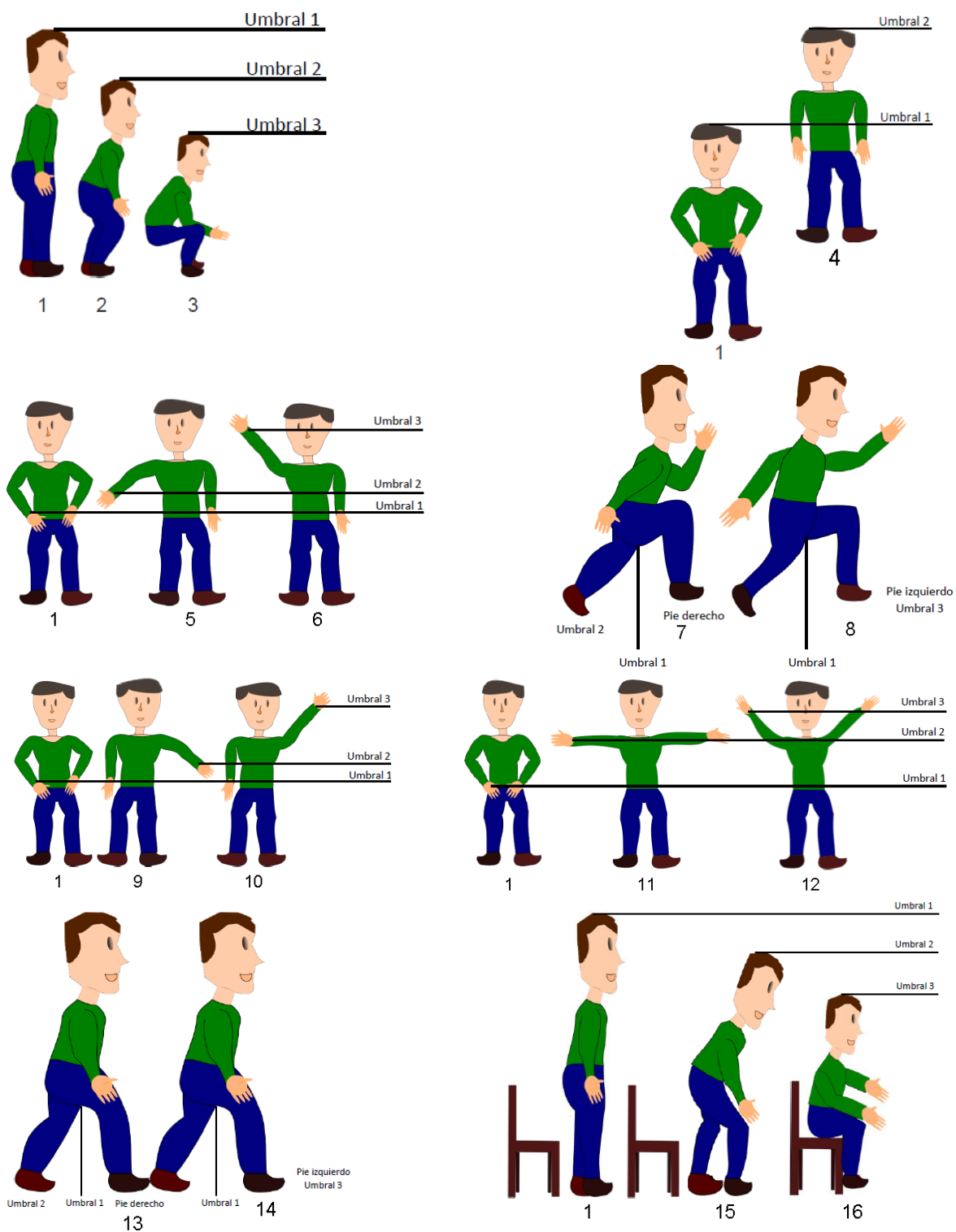
- Z_{n_D} Es la coordenada z normalizada del tobillo derecho.
- Z_{n_I} Es la Coordenada z normalizada del tobillo izquierdo.

■ Actividad 9 y 10

Se usa la metodología de la actividad 5 planteando tres umbrales que se muestran en la figura 8.

En la figura 8 se observan los umbrales de las etiquetas establecidas para cada una de las actividades de la tabla 1 teniendo en cuenta los movimientos primitivos definidos en la tabla 2. Se debe considerar que con estos umbrales no se podría realizar un algoritmo de decisión para la identificación de las actividades, ya que por sí solos no describen el comportamiento completo de cada una de las acciones. Como consecuencia se tendría que hacer un reconocimiento de posturas siendo este un enfoque diferente para el reconocimiento de actividades.

Figura 8. Umbrales para la asignación de etiquetas de los movimientos primitivos.



6. METODOLOGÍA

En la figura 9 se muestra la metodología propuesta para el reconocimiento de actividades físicas que se basa principalmente en el trabajo desarrollado por *Andrés et al* [13]. Este método realiza inicialmente la identificación de los movimientos primitivos para determinar la actividad realizada de acuerdo a la codificación o secuencia de los sub-movimientos. Este método nace de la relación que existe con el reconocimiento de voz. Considerando que una oración se compone de una secuencia de palabras aisladas o fonemas que de acuerdo a su codificación forman una frase [52, 20, 13]. A su vez, el modelo de reconocimiento de actividades parte de la idea de controlar una tarea a partir de la ejecución de una secuencia de sub-tareas más sencillas con el fin de controlar sistemas teleoperados remotamente por medio de un modelo híbrido entre los HMM y las SVM como lo propone [24]. De este modo, las SVM segmentan los datos de entrada y los HMM utilizan la salida de las SVM para determinar la secuencia más probable en consideración del tiempo. Como consecuencia se tiene un mejor desempeño tanto en el reconocimiento de los movimientos primitivos como de las actividades establecidas en la base de datos. Además, esta metodología permite ejecutar nuevas acciones en base al conjunto de movimientos conocidos. Esto es posible debido a las características de la máquina de vectores soporte para clasificar datos multidimensionales y a los HMM para trabajar con datos secuenciales con dependencia del tiempo [24, 53, 20]. Esta metodología presenta un mejor desempeño en comparación con otros enfoques donde sólo realizan la clasificación de las actividades físicas mediante las SVM o los HMM [20]. Por otro lado, es importante determinar la cantidad de movimientos primitivos para cada acción con el objetivo de que estos no se encuentren traslapados entre sí generando ambigüedades y dificultando el proceso de clasificación.

A continuación se definen las dos etapas para la identificación de actividades:

1. Reconocimiento de movimientos primitivos

- a) Codificación de movimientos primitivos.
- b) Extracción de características.

c) Clasificación de movimientos primitivos.

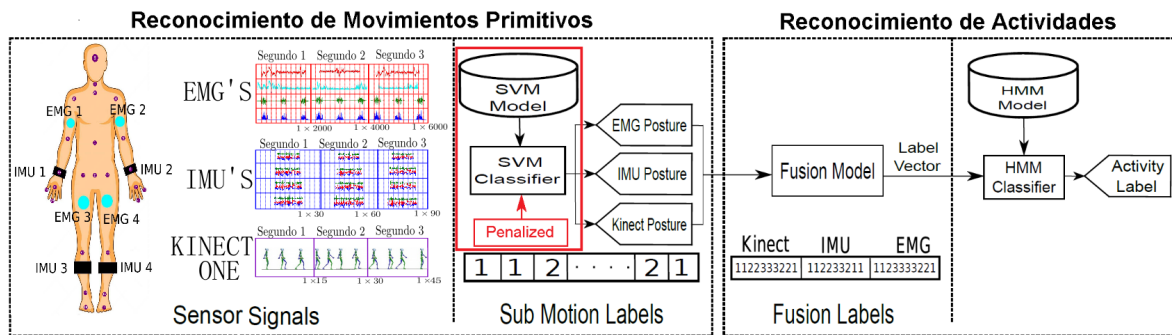
- Métodos de penalización.

2. Reconocimiento de actividades

a) Fusión de etiquetas de movimientos primitivos.

b) Clasificación de movimientos, entrenamiento y validación del modelo.

Figura 9. Metodología para el reconocimiento de actividades físicas humanas.



7. RECONOCIMIENTO DE MOVIMIENTOS PRIMITIVOS

Para el reconocimiento de los movimientos primitivos se han utilizado métodos supervisados y no supervisados usando una o dos modalidades de sensores para la captura del movimiento [21, 20, 23, 28, 12, 54], mientras que en [13, 14] utilizan tres modalidades de sensores como Kinect®, IMU y EMG para el análisis de movimientos. En este trabajo se propone realizar la identificación de los movimientos primitivos utilizando la base de datos anotada de la sección 5.2 y las SVM para penalizar los vectores de soporte por medio del parámetro de regularización C . Este método parte de la idea propuesta en [13, 23, 20] donde una actividad se compone de secuencia única de sub-movimientos para su clasificación y que las SVM presentan una buena eficiencia en muchos problemas de clasificación multidimensional al momento de encontrar un hiperplano óptimo de separación entre las clases. Teniendo en cuenta que para un problema multiclase la estrategia que se prefiere de la SVM es uno contra uno (one vs one), ya que presenta mejores resultados en comparación con el enfoque de uno contra todos (one vs all) como lo menciona [23] y [20].

7.1. Codificación de movimientos primitivos

Para esta sección se diseñó un vocabulario o diccionario que describe el conjunto de sub-movimientos necesarios para realizar las actividades propuestas en la tabla 1. En las tablas 2 y 3 se muestran la lista de movimientos primitivos junto con las respectivas etiquetas asociadas a cada actividad. En las figuras 10, 11, 12, 13 y 14 se observa la ejecución de cada una de las actividades.

Tabla 3. Etiquetas de la SVM para cada postura

| Sub-movimiento | Etiqueta | Sub-movimiento | Etiqueta |
|--|----------|--|----------|
| Reposo | 1 | Mano izquierda levantada a $\frac{1}{4}$ | 9 |
| Medio Agachado | 2 | Mano izquierda levantada a $\frac{3}{4}$ | 10 |
| Totalmente agachado | 3 | Manos levantadas a $\frac{1}{4}$ | 11 |
| Suspendido en el aire $\frac{1}{4}$ | 4 | Manos levantadas a $\frac{3}{4}$ | 12 |
| Mano derecha levantada a $\frac{1}{4}$ | 5 | Paso adelante pie derecho | 13 |
| Mano derecha levantada a $\frac{3}{4}$ | 6 | Paso adelante pie izquierdo | 14 |
| Pierna derecha hacia adelante con rodilla flexionada | 7 | Medio sentado | 15 |
| Pierna izquierda hacia adelante con rodilla flexionada | 8 | Sentado | 16 |

Figura 10. Actividades 1 y 3.

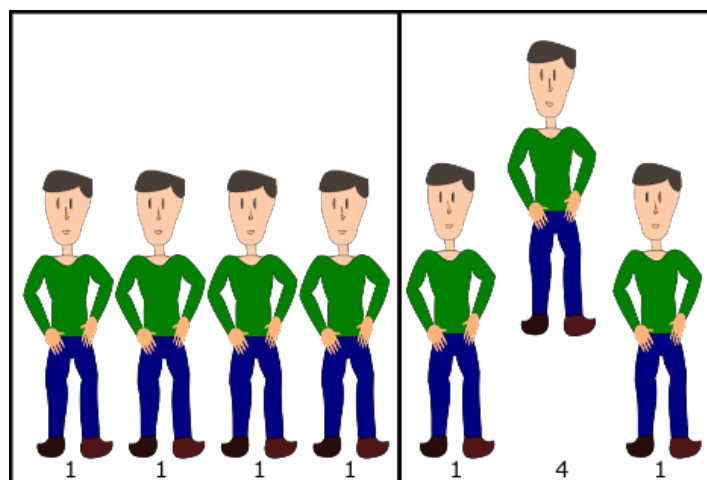


Figura 11. Actividades 2 y 8.

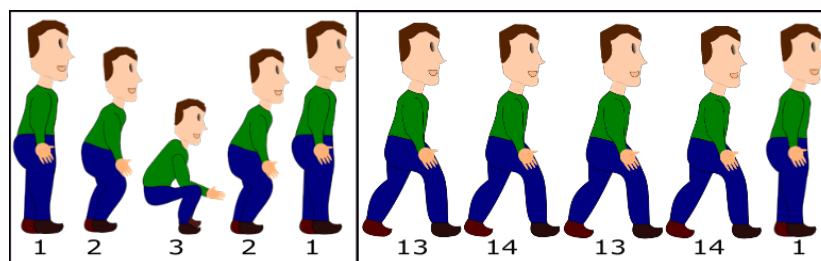


Figura 12. Actividades 4, 6 y 7.

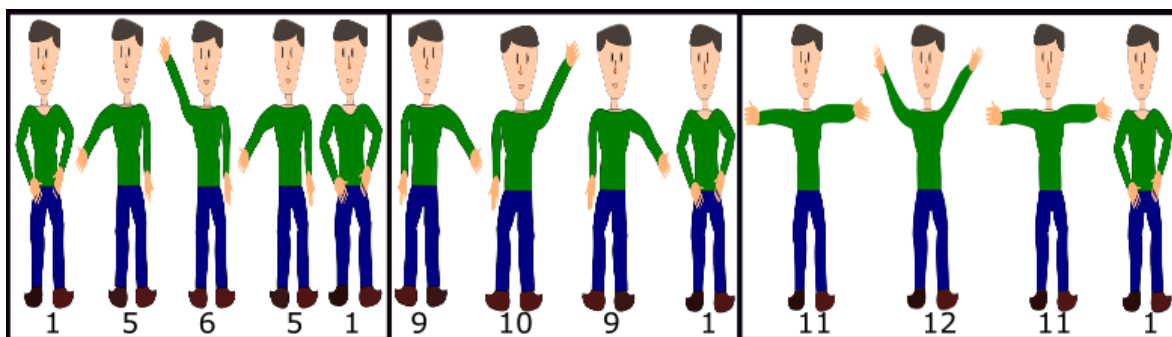


Figura 13. Actividad 5.

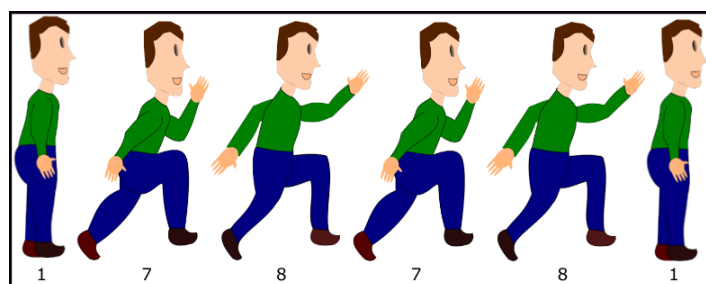
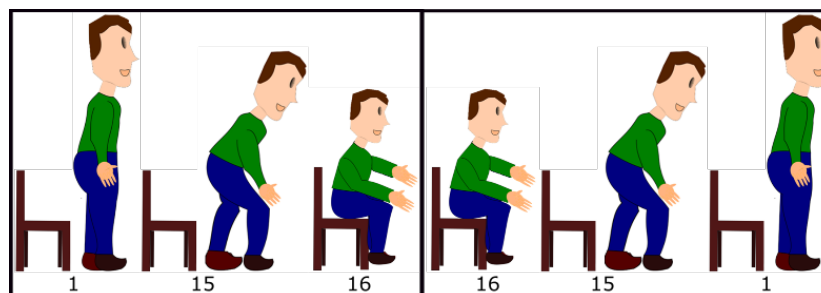


Figura 14. Actividades 9 y 10.



7.1.1. Tamaños de ventana

Para realizar la extracción de características primero se determina una tamaño de ventana de acuerdo a la modalidad del sensor y su frecuencia de muestreo. En este caso se decide utilizar dos métodos:

1. Ventanas definidas por *Andrés et al* [13]

Análisis de forma experimental observando gráficamente la dinámica de los datos para cada movimiento primitivo.

2. Autocorrelación de los datos

Reducción de dimensionalidad de los datos mediante el análisis de las componentes principales (PCA) para auto-correlacionar la información de las señales provistas por los sensores. De esta manera se determina la frecuencia de cada sub-movimiento contando el número de muestras cuando se detectan los cruces por ceros que corresponden al tamaño de la ventana.

7.1.2. Extracción de características del Kinect One®

Para este caso, se utilizan los 25 puntos articulados del cuerpo provistos por el Kinect One® como se observa en la figura 1 adquiridos a una frecuencia de 15 muestras por segundo [44, 55].

La caracterización de estos datos consiste en el cálculo de los siguientes descriptores:

- Características Polares [56].
- Descriptores estadísticos.
- Ángulos de movimiento del miembro superior e inferior.

El cálculo de las características polares se basa en la transformación de los puntos articulados de coordenadas cartesianas J a coordenadas polares cilíndricas P . El procedimiento para calcular el descriptor es el siguiente:

Dado:

$$J = [X_1 Y_1 Z_1 \dots X_{25} Y_{25} Z_{25}] \quad (39)$$

1. Se calcula el centro de masa del objetivo evaluado.

2. Se transforma cada articulación a coordenadas polares utilizando como origen de coordenadas el centro de masa, lo que permite obtener el vector.

$$P_i = [r_1 \theta_1 r_2 \theta_2 \dots r_{25} \theta_{25}] \quad (40)$$

Donde i es la i -ésima muestra de la ventana de tres muestras, $i = \{1, 2, 3\}$.

Para el cálculo de los descriptores estadísticos, se utilizan las fórmulas de la tabla 4.

Tabla 4. Descriptores estadísticos utilizados.

| Nombre | Descripción | Nombre | Descripción |
|-------------------|--|----------------------|---|
| Media de x | $m_x = \frac{1}{3} \sum_{i=1}^3 x_i$ | Varianza de x | $v_x = \frac{1}{3} \sum_{i=1}^3 (x_i - \bar{x})^2$ |
| Media de y | $m_y = \frac{1}{3} \sum_{i=1}^3 y_i$ | Varianza de y | $v_y = \frac{1}{3} \sum_{i=1}^3 (y_i - \bar{y})^2$ |
| Media de z | $m_z = \frac{1}{3} \sum_{i=1}^3 z_i$ | Varianza de z | $v_z = \frac{1}{3} \sum_{i=1}^3 (z_i - \bar{z})^2$ |
| Media de r | $m_r = \frac{1}{3} \sum_{i=1}^3 r_i$ | Varianza de r | $v_r = \frac{1}{3} \sum_{i=1}^3 (r_i - \bar{r})^2$ |
| Media de θ | $m_\theta = \frac{1}{3} \sum_{i=1}^3 \theta_i$ | Varianza de θ | $v_\theta = \frac{1}{3} \sum_{i=1}^3 (\theta_i - \bar{\theta})^2$ |

Los ángulos de movimiento del miembro superior e inferior se basan en la flexión y extensión del codo y la rodilla respectivamente con los puntos articulados del hombro, codo y muñeca para el miembro superior y los puntos articulados de la cadera, rodilla y tobillo para el miembro inferior como se observa en la tabla 5.

Para determinar estos ángulos se establecen los siguientes pasos:

- Determinar los vectores para el miembro superior formados entre los puntos:

- P_{Hj} y P_{Cj} (**V1j**)

- P_{Mj} y P_{Cj} (**V2j**)

- Determinar los vectores para el miembro inferior formados entre los puntos:

- P_{CAj} y P_{Rj} (**V2j**)

Tabla 5. Puntos articulados del Kinect One®.

| Miembro | Articulación | Punto |
|----------|------------------------|--------------------------------------|
| Superior | Hombro derecho (HD) | $P_{HD} = (X_8, Y_8, Z_8)$ |
| Superior | Hombro izquierdo (HI) | $P_{HI} = (X_4, Y_4, Z_4)$ |
| Superior | Codo derecho (CD) | $P_{CD} = (X_9, Y_9, Z_9)$ |
| Superior | Codo izquierdo (CI) | $P_{CI} = (X_5, Y_5, Z_5)$ |
| Superior | Muñeca derecha (MD) | $P_{MD} = (X_{10}, Y_{10}, Z_{10})$ |
| Superior | Muñeca izquierda (MI) | $P_{MI} = (X_6, Y_6, Z_6)$ |
| Inferior | Cadera derecha (CAD) | $P_{CAD} = (X_{16}, Y_{16}, Z_{16})$ |
| Inferior | Cadera izquierda (CAI) | $P_{CAI} = (X_{12}, Y_{12}, Z_{12})$ |
| Inferior | Rodilla derecha (RD) | $P_{RD} = (X_{17}, Y_{17}, Z_{17})$ |
| Inferior | Rodilla izquierda (RI) | $P_{RI} = (X_{13}, Y_{13}, Z_{13})$ |
| Inferior | Tobillo derecho (TD) | $P_{TD} = (X_{18}, Y_{18}, Z_{18})$ |
| Inferior | Tobillo izquierdo (TI) | $P_{TI} = (X_{14}, Y_{14}, Z_{14})$ |

- P_{Tj} y P_{Rj} (**V2j**)

Donde j corresponde al punto articulado izquierdo o derecho.

- Calcular los ángulos de flexión y extensión del codo y de la rodilla (θ_{cj} , θ_{rj}) usando la ecuación 35 para el lado derecho e izquierdo.

Por lo tanto, para un ventana de observación de k muestras el vector de características KIT se define como:

$$\text{KIT} = [P_1 P_2 P_3 m_x m_y m_z m_r m_\theta v_x v_y v_z v_r v_\theta \theta_{cD} \theta_{cI} \theta_{rI} \theta_{rD}]_{1 \times 89} \quad (41)$$

7.1.3. Extracción de características para red sensores IMU

Teniendo en cuenta la información por los 4 acelerómetros tri-axiales (a_x, a_y, a_z) con tres grados de libertad distribuidos en diferentes partes del cuerpo (ver figura 1) se calcularon las orientaciones Roll y Pitch realizando la conversión a coordenadas esféricas por medio de las siguientes ecuaciones:

Fuerza gravitacional G .

$$G = Raw_{accel} \left[\frac{R}{2^n - 1} \right] \quad (42)$$

Donde:

- n es la resolución en bits del conversor análogo digital.
- R es el rango de del voltaje de entrada del conversor A/D.
- Raw_{accel} es la señal de aceleración del sensor.

$$Pitch = \arctan \left(\frac{G_y}{\sqrt{G_x^2 + G_z^2}} \right) \quad (43)$$

$$roll = \arctan \left(\frac{-G_x}{G_z} \right) \quad (44)$$

Construyendo el vector $I_k = \begin{bmatrix} a_x & a_y & a_z & Pitch & Roll \end{bmatrix}_{1 \times 5}$ para cada instante de tiempo. Donde k es el k -ésimo acelerómetro con $k = \{1, 2, 3, 4\}$. La frecuencia de muestreo para cada sensor fue de $30Hz$. La caracterización del vector I se calculan los siguientes descriptores:

- Características basadas en los parámetros físicos de movimiento humano [57].
- Descriptores estadísticos

Para determinar las características basadas en los parámetros físicos del movimiento humano se utilizaron las fórmulas de la tabla 6.

Tabla 6. Características IMU.

| Nombre | Definición |
|---|---|
| <i>Mean of Movement Intensity</i> (AI) | $AI = \frac{1}{3} \sum_{i=1}^3 MI(i),$ $MI = \ A_{xyz}\ _2$ $A_{xyz} = \begin{bmatrix} a_x & a_y & a_z \end{bmatrix}^T$ |
| <i>Variance of MI</i> (VI) | $VI = \frac{1}{3} \sum_{i=1}^3 (MI(i) - AI)^2$ |
| <i>Normalized Signal Magnitude Area</i> (SMA) | $SMA = \frac{1}{3} \left(\sum_{i=1}^3 a_x(i) + \sum_{i=1}^3 a_y(i) + \sum_{i=1}^3 a_z(i) \right)$ |
| <i>Eigenvalues of Dominant Directions</i> (EVA) | Vectores propios de la matriz de covarianza de la aceleración a lo largo de los ejes x,y,z en cada ventana de observación |
| <i>Averaged Acceleration Energy</i> (AAE) | Valor medio de la energía a lo largo de tres ejes de aceleración donde la energía es la suma de las magnitudes de los componentes discretos FFT al cuadrado de la señal de cada eje del sensor, normalizado por la longitud de la ventana |
| <i>Averaged Rotation Energy</i> (ARE) | Valor medio de la energía a lo largo de los ángulos de rotación, para este caso solo Roll y Pitch |

Para el cálculo de los descriptores estadísticos se utilizan las fórmulas de la tabla 7.

Tabla 7. Descriptores estadísticos utilizados

| Nombre | Descripción | Nombre | Descripción |
|----------------|---|----------------------|--|
| Media de a_x | $m_{ax} = \frac{1}{N} \sum_{i=1}^N ax_i$ | Varianza de x | $v_{ax} = \frac{1}{N} \sum_{i=1}^N (ax_i - \bar{a}_x)^2$ |
| Media de a_y | $m_{ay} = \frac{1}{N} \sum_{i=1}^N ay_i$ | Varianza de y | $v_{ay} = \frac{1}{N} \sum_{i=1}^N (ay_i - \bar{a}_y)^2$ |
| Media de a_z | $m_{az} = \frac{1}{N} \sum_{i=1}^N az_i$ | Varianza de z | $v_{az} = \frac{1}{N} \sum_{i=1}^N (az_i - \bar{a}_z)^2$ |
| Media de a_r | $m_{ar} = \frac{1}{N} \sum_{i=1}^N ar_i$ | Varianza de r | $v_{ar} = \frac{1}{N} \sum_{i=1}^N (ar_i - \bar{a}_r)^2$ |
| Media de a_p | $m_{a\theta} = \frac{1}{N} \sum_{i=1}^N ap_i$ | Varianza de θ | $v_{ap} = \frac{1}{N} \sum_{i=1}^N (ap_i - \bar{a}_p)^2$ |

Aplicando los cálculos de las tablas 6 y 7, se obtiene el vector de características IM en cada ventana de 3 muestras.

$$IM = [AI \ VI \ SMA \ EVA \ AAE \ ARE \ m_{ax} \ m_{ay} \ m_{az} \ m_{ar} \ m_{ap} \ v_{ax} \ v_{ay} \ v_{ar} \ v_{ap}]_{1 \times 15} \quad (45)$$

7.1.4. Extracción de características Red de sensores EMG

La frecuencia de muestreo para cada sensor electromiográfico es de $2KHz$ como se observa en el figura 1, obteniendo una señal análoga E_i , donde i es el i -ésimo sensor EMG. Cada ventana V_j es caracterizada por medio de una transformada Waveleth, obteniendo un vector de características M_i de 1×2000 , donde i es el i -ésimo sensor. Entonces, para cada ventana j se obtendrá el vector de características.

$$EMG_j = \begin{bmatrix} M_1 & M_2 & M_3 & M_4 \end{bmatrix}_{1 \times 2000}$$

7.2. Clasificación de movimientos primitivos

Para esta etapa se utilizaron tres máquinas de vectores de soporte multiclase con estrategia de clasificación one vs all para el método de penalización C-SVC y one vs one para el segundo método SVM Binarias Ponderadas. Para ambos métodos se estableció un kernel gaussiano con radio 1×10^4 usando la base de datos de la sección 5.2. En este caso, se desea penalizar las SVM teniendo en cuenta algunos de los problemas que se encontraron en [58] y [13], como un alto traslape y desbalanceo de las clases, además de las oclusiones parciales o auto-oclusiones del Kinect® o pérdidas de conexión en los sistemas de adquisición de las señales de los IMU's o EMG's bajando la eficiencia del clasificador en la identificación de los movimientos primitivos.

Los métodos propuestos mencionados en las secciones 4.1.3 y 4.1.4 pretermiten penalizar el parámetro de regularización C de las SVM considerando el tamaño de los datos del conjunto de entrenamiento. El valor de este parámetro puede ser muy pequeño ($C > 0$) o muy grande ($C \rightarrow \infty$) permitiendo la penalización de los puntos que se encuentran en el lado equivocado del límite del margen y minimizando de esta manera el error de entrenamiento. Considerando que la variable slack ξ_i permite relajar el margen del clasificador, este es sensible a datos que sean atípicos aumentando el error de forma lineal. Por lo tanto, es importante elegir un valor adecuado para C de manera eficiente. Algunos autores recomiendan realizar una grilla de observación encontrando un valor para C dentro del rango, esto con el objetivo de maximizar

el margen mientras se penalizan los puntos que están en el lado equivocado de este [59, 6]. Por lo cual, se escogió una grilla de observación de $[1 \times 10^{-3}, 1 \times 10^3]$ para encontrar un valor C para cada clase.

Considerando que ambas metodologías tienen en cuenta el tamaño del conjunto de datos de entrenamiento en la penalización del parámetro de regularización C , en cada caso se realiza de forma diferente. Partiendo que las SVM es un método biclase, con C-SVC se tiene un valor de C donde se penaliza con respecto al tamaño del conjunto de datos a clasificar; mientras que con las SVM binaria ponderada se tienen dos valores de C para cada una de las clases penalizando de acuerdo al tamaño de los datos de cada clase.

8. RECONOCIMIENTO DE ACTIVIDADES

Teniendo en cuenta la lista de actividades de la tabla 1 y la información de la máquina de vectores de soporte correspondiente a las secuencias de los movimientos primitivos para cada modalidad de sensor, se efectúa la fusión de estos datos para clasificarlos por medio de los modelos ocultos de Markov (HMM) y de esta manera identificar la actividad física de acuerdo a la secuencia de los sub-movimientos como se observa en la figura 9.

8.1. Fusión de datos

Para la fusión de los datos correspondientes a las etiquetas entregadas por las SVM penalizadas se crea un vector de características EF que concatena linealmente y sin ponderación las etiquetas generadas por los clasificadores de cada sensor durante la ventana de observación de tres segundos, como se observa en la ecuación 46.

$$EF = [[EK_1 EK_2 \dots EK_{15}] [EI_1 EI_2 \dots EI_{30}] [EE_1 EE_2 \dots EE_{60}]]_{105 \times 1} \quad (46)$$

Donde:

- EK : Vector de etiquetas de la SVM penalizada provista por el Kinect One®.
- EI : Vector de etiquetas de la SVM penalizada provista por la red de sensores IMU.
- EE : Vector de etiquetas de la SVM penalizada provista por la red de sensores EMG.

8.2. Clasificación, entrenamiento y validación del modelo

Teniendo en cuenta que este proceso de clasificación parte de un modelo entrenado por las SVM el cual proporciona una separación entre cada clase, en esta etapa se realiza la evaluación del modelo utilizando la estrategia de validación cruzada mediante las iteraciones de Montecarlo con criterio de paro $\|\text{diag}(M_k) - \text{diag}(M_{k-1})\| < 0,001$. Donde $\text{diag}(M_k)$ es el vector generado por la diagonal de la matriz de confusión y k es la iteración promedio actual del Montecarlo. El conjunto de datos de entrenamiento para este caso fue del 70 % y para los

de prueba del 30% con la base de datos de la sección 5.1 donde la fragmentación de esta se realiza de forma aleatoria para cada iteración. Lo cual permite observar el comportamiento de los HMM en la clasificación de la actividad con diferentes datos de entrada. Para los HMM se usaron 24 estados y 32 centroides en la construcción del libro de códigos.

El método de paro del Montecarlo consiste en encontrar un desempeño mayor al 95% o hasta que se cumplan 100 iteraciones escogiendo el modelo con mejor desempeño. Para evaluar el desempeño del clasificador se calcula la matriz de confusión y se evalúa el número total de aciertos por iteración ($Promedio_{MC}$), determinado el comportamiento promedio ($Desempeño$) e incertidumbre ($\pm\Delta D$) de todos los resultados obtenidos.

Algoritmo:

```

for  $k = 1 : n$ 
     $Promedio_{MC}[k] = mean(diag(M_k))$ 
end
 $Desempeño = mean(Promedio_{MC})$ 
 $\pm\Delta D = std(Promedio_{MC})$ 

```

9. EXPERIMENTOS Y RESULTADOS

En esta sección se muestran los resultados experimentales que se dividen en dos partes. Primero se realiza un estudio comparativo de dos métodos de penalización de las SVM (C-SVM y SVM binaria ponderada) con respecto al modelo propuesto en *Andrés et al* [13] considerando la variación del tamaño de ventana para la identificación de los movimientos primitivos. La metodología de *Andrés et al* [13] utiliza una máquina de vectores de soporte multiclase con la estrategia de clasificación one vs all y un kernel gaussiano para las tres modalidades de sensores (Kinect One®, IMU's y EMG's). Segundo, se realiza la clasificación de las actividades físicas por medio de los HMM evaluando la eficiencia del clasificador con una modalidad de sensor y la fusión de dos y tres tipos de sensores como lo propone *Andrés et al* [13] con la siguiente combinación:

- Kinect One® + IMU
- Kinect One® + EMG
- IMU + EMG
- Kinect One® + IMU + EMG

Para cada caso se validó el desempeño de los clasificadores mediante el experimento Montecarlo para la estimación del error. Es de aclarar, que todos los resultados experimentales del metodología sólo muestran la diagonal de las matrices de confusión, ya que se evidencia una alta tasa de desempeño con poco traslape entre clases. Sin embargo, en los anexos se incluirán las matrices de confusión de todos los experimentos realizados.

Para la identificación de cada resultado se tienen los siguiente parámetros:

1. Selección del tamaño de ventana

- V_CyH: Tamaño definido por *Andrés et al* [13].

Este análisis es de forma experimental donde se observa gráficamente la dinámica de los datos para cada movimiento primitivo.

- V_A-Corr: Determinación del tamaño por medio de la auto-correlación de los datos.

Se realiza la reducción de dimensionalidad de los datos mediante el análisis de las componentes principales (PCA) para auto-correlacionar la información con las características principales de las señales provistas por los sensores. De esta manera se determina la frecuencia de cada sub-movimiento contando el número de muestras cuando se detectan los cruces por ceros los cuales corresponden al tamaño de la ventana. Esto es posible ya que la información sensor corresponde a la realización de actividades físicas que tienen por lo menos uno o más ciclos de los movimientos primitivos como se observa en las figuras 10, 11, 12, 13 y 14.

2. Experimento para el análisis de movimientos primitivos

- SVM_CyH: SVM sin penalizar propuesto por *Andrés et al* [13].
- C-SVM: Método 1 de penalización de la SVM.
- SVM_BP: Método 2 de penalización de la SVM.
- C-SVM_V_A-Corr: Método 1 de penalización de las SVM con el tamaño de ventana obtenido en V_A-Corr.

3. Experimento de reconocimiento de actividades físicas

- HMM_CyH: Clasificador HMM con las etiquetas entregadas por SVM_CyH.
- HMM_C: Clasificador HMM con las etiquetas entregadas por C-SVM.
- HMM_BP: Clasificador HMM con las etiquetas entregadas por SVM_BP.

Se aplica la metodología diseñada con las siguientes combinaciones de sensores:

- Kinect One®
- IMUs
- EMGs
- Kinect One® + IMUs
- Kinect One® +EMGs

- IMUs+EMGs
- Kinect One® + IMUs + EMGs

9.1. Tamaños de ventanas

Considerando las frecuencias de muestreo para cada uno de los sensores, se establecen los tamaños de ventanas definidos en el trabajo de *Andrés et al* [13] y los calculados con la autocorrelación de los datos.

1. Tamaños de ventanas definidos en *Andrés et al* [13]

- Para el Kinect One®: El tamaño de ventana es de 3 muestras. En consecuencia, se obtiene información a una frecuencia de $5Hz$.
- Para la redes de sensores IMU: El tamaño de ventana es de 3 muestras. Por lo cual, se obtiene información a una frecuencia de $10Hz$.
- Para la redes de sensores EMGs: El tamaño de ventana es es de 100 muestras. Por lo tanto se obtiene información a una frecuencia de $20Hz$.

1. Tamaños de ventanas definidos con la autocorrelación de datos

- Para el Kinect One®: El tamaño de ventana es de 5 muestras. Por lo cual, se obtiene información a una frecuencia de $3Hz$.
- Para la redes de sensores IMU: El tamaño de ventana es igual al determinado *Andrés et al* [13].
- Para la redes de sensores EMGs: El tamaño de ventana es es de 8 muestras. Por lo tanto, se obtiene información a una frecuencia de $250Hz$.

En la tabla 8 se visualizan cada uno de los resultados.

Tabla 8. Tamaños de ventanas.

| Sensor | Frecuencia de muestreo [muestras/s] | Ventana | | | |
|-------------|--|-----------------|--------------------|-----------------|--------------------|
| | | V_CyH | | V_A-Corr | |
| | | <i>Muestras</i> | <i>#Ventanas/s</i> | <i>Muestras</i> | <i>#Ventanas/s</i> |
| Kinect One® | 15 | 3 | 5 | 5 | 3 |
| IMU | 30 | 3 | 10 | 3 | 10 |
| EMG | 2000 | 100 | 20 | 8 | 250 |

9.2. Análisis de movimientos primitivos

En la tabla 9 se observan los promedios de las trazas de las matrices de confusión de cada modelo por medio del experimento de Montecarlo descrito en la sesión 8.2. En este sentido, el Kincet One® con cualquiera de los modelos de aprendizaje de máquina especificados con anterioridad presenta eficiencias mayores al 90 % en la clasificación de los movimientos primitivos. Por otro lado, los dos métodos de penalización presentan un mejor desempeño con respecto *Andrés et al* [13]. Además, el método 2 (SVM binarias ponderada) presenta la mayor eficiencia del $95,15 \% \pm 0,00 \%$ con los otros experimentos. Sin embargo, la clase 16 tiene un bajo rendimiento del $23,34 \% \pm 0,00 \%$ debido al traslape las demás clases como evidencia en los anexos, mientras que con el método 1 (C-SVM) no se tiene este comportamiento obteniendo una tasa de acierto del $92,68 \% \pm 0,40 \%$.

Los resultados obtenidos de los diferentes modelos de aprendizaje utilizando la red de sensores IMU muestran que el método 2 presenta un mejor rendimiento en comparación con las otras variaciones del las SVM y tamaños de ventana. Además, este valor es superior al resultado del Kinect One®. En este caso el porcentaje de acierto con esta modalidad de sensor es del $96,86 \% \pm 0,00 \%$.

Tabla 9. Traza de las matrices de confusión con Kinect One®.

| Clases | V_A-Corr | V_CyH | | |
|----------|-------------------|------------------|------------------|-------------------|
| | C-SVM_V_A-Corr | SVM_CyH | C-SVM | SVM_BP |
| 1 | 83,68 \pm 0,59 | 92,59 \pm 2,24 | 93.71 \pm 0.66 | 100.00 \pm 0.00 |
| 2 | 94,94 \pm 1,45 | 87.29 \pm 3.07 | 87.89 \pm 0.87 | 100.00 \pm 0.00 |
| 3 | 86,45 \pm 1,75 | 85.93 \pm 3.91 | 87.00 \pm 0.17 | 100.00 \pm 0.00 |
| 4 | 80,51 \pm 3,28 | 95.96 \pm 3.63 | 93.36 \pm 0.13 | 100.00 \pm 0.00 |
| 5 | 93,00 \pm 1,16 | 87.41 \pm 4.10 | 96.69 \pm 3.51 | 100.00 \pm 0.00 |
| 6 | 96,37 \pm 1,10 | 95.31 \pm 2.10 | 93.50 \pm 1.09 | 100.00 \pm 0.00 |
| 7 | 88,72 \pm 1,83 | 88.88 \pm 3.66 | 89.30 \pm 0.41 | 100.00 \pm 0.00 |
| 8 | 93,76 \pm 1,37 | 89.12 \pm 3.13 | 95.83 \pm 0.26 | 100.00 \pm 0.00 |
| 9 | 88,65 \pm 1,97 | 92.55 \pm 2.90 | 87.69 \pm 0.27 | 100.00 \pm 0.00 |
| 10 | 93,90 \pm 0,69 | 95.01 \pm 2.69 | 93.70 \pm 0.84 | 100.00 \pm 0.00 |
| 11 | 99,73 \pm 1,18 | 92.05 \pm 3.47 | 88.38 \pm 1.67 | 100.00 \pm 0.00 |
| 12 | 100,00 \pm 0,00 | 96.57 \pm 2.29 | 98.86 \pm 0.26 | 100.00 \pm 0.00 |
| 13 | 90,71 \pm 0,87 | 93.78 \pm 2.73 | 96.06 \pm 0.26 | 100.00 \pm 0.00 |
| 14 | 88,41 \pm 0,73 | 92.55 \pm 3.47 | 93.36 \pm 0.84 | 100.00 \pm 0.00 |
| 15 | 84,99 \pm 0,93 | 89.99 \pm 2.80 | 94.91 \pm 1.67 | 100.00 \pm 0.00 |
| 16 | 91,96 \pm 0,58 | 90.64 \pm 1.60 | 92.64 \pm 0.16 | 23.34 \pm 0.00 |
| Promedio | 90,99 \pm 0,41 | 91,60 \pm 0,91 | 92.68 \pm 0.40 | 95.15 \pm 0.00 |

Tabla 10. Diagonal de las matrices de confusión con la red de sensores IMU.

| Clases | V_A-Corr | V_CyH | | |
|----------|------------------|------------------|-------------------|-------------------|
| | C-SVM_V_A-Corr | SVM_CyH | C-SVM | SVM_BP |
| 1 | 85.54 \pm 0.09 | 87.40 \pm 0.52 | 90.14 \pm 0.00 | 61.85 \pm 0.000 |
| 2 | 63.77 \pm 0.84 | 62.37 \pm 2.62 | 64.42 \pm 0.00 | 100.00 \pm 0.00 |
| 3 | 65.49 \pm 0.63 | 66.72 \pm 3.47 | 65.26 \pm 0.000 | 100.00 \pm 0.00 |
| 4 | 69.97 \pm 1.73 | 77.91 \pm 5.51 | 91.67 \pm 0.00 | 100.00 \pm 0.00 |
| 5 | 69.73 \pm 1.02 | 55.97 \pm 7.32 | 73.08 \pm 0.00 | 100.00 \pm 0.00 |
| 6 | 81.35 \pm 0.52 | 80.02 \pm 2.17 | 74.86 \pm 0.00 | 100.00 \pm 0.00 |
| 7 | 76.25 \pm 0.41 | 70.57 \pm 3.58 | 75.00 \pm 0.00 | 100.00 \pm 0.00 |
| 8 | 70.81 \pm 1.93 | 72.65 \pm 3.35 | 79.46 \pm 0.00 | 100.00 \pm 0.00 |
| 9 | 68.88 \pm 1.15 | 62.94 \pm 5.16 | 69.51 \pm 0.00 | 100.00 \pm 0.00 |
| 10 | 86.35 \pm 0.37 | 62.94 \pm 5.16 | 85.80 \pm 0.00 | 100.00 \pm 0.00 |
| 11 | 71.96 \pm 0.87 | 64.80 \pm 5.97 | 71.43 \pm 0.00 | 100.00 \pm 0.00 |
| 12 | 84.11 \pm 0.71 | 85.61 \pm 1.83 | 85.87 \pm 0.00 | 100.00 \pm 0.00 |
| 13 | 74.84 \pm 0.92 | 74.16 \pm 2.93 | 77.14 \pm 0.00 | 97.30 \pm 0.00 |
| 14 | 70.16 \pm 0.46 | 73.85 \pm 2.72 | 74.21 \pm 0.00 | 90.57 \pm 0.00 |
| 15 | 77.27 \pm 0.64 | 73.45 \pm 3.32 | 72.37 \pm 0.00 | 100.00 \pm 0.00 |
| 16 | 88.54 \pm 0.18 | 88.83 \pm 1.14 | 88.87 \pm 0.00 | 100.00 \pm 0.00 |
| Promedio | 75.31 \pm 0.22 | 73.71 \pm 1.03 | 77.44 \pm 0.00 | 96.86 \pm 0.00 |

En el análisis del desempeño de los datos provistos por la red de sensores EMG presentan una baja confiabilidad en comparación con las otras modalidades de sensores. En este caso, el método 2 (SVM_BP) tiene un desempeño del 67,51 % y la clase 8 presentó una bajo porcentaje de acierto tanto para SVM_BP como para SVM_CyH, mientras que con C-SVM la clase 4 tiene un desempeño menor al 30 % que es comparable con la clase 1 del método 2 de penalización para el reconocimiento de los movimientos primitivos.

Por otro lado, se evidencia que el segundo método de penalización presenta un mayor desempeño con respecto a los otros enfoques en todas las modalidades de sensores como lo muestran las tablas 9, 10 y 11. Además, los tamaños de ventanas determinados con la autocorrelación de los datos para las diferentes modalidades de sensores tienen un menor rendimiento en comparación con las ventanas definidas por *Andrés et al* [13].

Tabla 11. Diagonal de las matrices de confusión con la red de sensores EMG.

| Clases | V_A-Corr | V_CyH | | |
|----------|------------------|-------------------|------------------|------------------|
| | C-SVM_V_A-Corr | SVM_CyH | C-SVM | SVM_BP |
| 1 | 1,30 \pm 1,30 | 56,47 \pm 11,68 | 50,00 \pm 0,00 | 22,22 \pm 0,00 |
| 2 | 26,84 \pm 0,37 | 51,86 \pm 2,10 | 47,80 \pm 0,00 | 44,82 \pm 0,00 |
| 3 | 33,77 \pm 0,48 | 57,89 \pm 4,18 | 65,31 \pm 0,00 | 83,87 \pm 0,00 |
| 4 | 18,53 \pm 0,09 | 72,02 \pm 6,81 | 28,45 \pm 0,00 | 71,43 \pm 0,00 |
| 5 | 37,65 \pm 0,29 | 68,54 \pm 4,07 | 71,88 \pm 0,00 | 72,73 \pm 0,00 |
| 6 | 47,66 \pm 0,66 | 68,73 \pm 1,82 | 74,69 \pm 0,00 | 78,87 \pm 0,00 |
| 7 | 28,96 \pm 0,29 | 49,04 \pm 3,36 | 51,85 \pm 0,00 | 82,89 \pm 0,00 |
| 8 | 32,18 \pm 1,31 | 44,31 \pm 2,86 | 63,78 \pm 0,00 | 11,85 \pm 0,00 |
| 9 | 34,52 \pm 1,25 | 70,46 \pm 4,52 | 59,09 \pm 0,00 | 78,95 \pm 0,00 |
| 10 | 53,10 \pm 0,42 | 68,06 \pm 2,78 | 82,40 \pm 0,00 | 87,17 \pm 0,00 |
| 11 | 28,05 \pm 0,11 | 65,88 \pm 4,60 | 58,06 \pm 0,00 | 81,37 \pm 0,00 |
| 12 | 31,48 \pm 0,46 | 64,80 \pm 2,69 | 67,02 \pm 0,00 | 84,61 \pm 0,00 |
| 13 | 37,17 \pm 0,57 | 57,35 \pm 3,19 | 66,67 \pm 0,00 | 68,46 \pm 0,00 |
| 14 | 49,67 \pm 0,03 | 52,71 \pm 2,54 | 70,56 \pm 0,00 | 69,90 \pm 0,00 |
| 15 | 40,58 \pm 0,88 | 53,01 \pm 2,85 | 64,83 \pm 0,00 | 66,67 \pm 0,00 |
| 16 | 52,54 \pm 0,01 | 56,65 \pm 0,90 | 52,61 \pm 0,00 | 74,42 \pm 0,00 |
| Promedio | 34,63 \pm 0,29 | 59,86 \pm 1,29 | 60,94 \pm 0,00 | 67,51 \pm 0,00 |

En la tabla 12 se muestran los resultados de cada metodología en orden mayor a menor desempeño para cada modalidad de sensor. Se identifica que los dos métodos de penalización de

las SVM presentan un mejor desempeño para las tres modalidades de sensores en comparación con SVM_CyH, siendo el segundo método (SVM_BP) en obtener eficiencias superiores al 95 % con el Kinect One® y los IMUs, para el caso de los sensores EMGs se tienen resultados comparables con los tres experimentos. También, se muestra que para el caso de los sensores EMG una ventana pequeña con una alta frecuencia de muestreo en la adquisición de los datos no es recomendable, ya que presenta una eficiencia inferior en comparación con la ventana definida por *Andrés et al* [13].

Tabla 12. Comportamiento promedio de la metodología para el reconocimiento de los movimientos primitivos.

| Puesto | Kinect One® | | IMUs | | EMGs | |
|--------|----------------|------------------|----------------|------------------|----------------|------------------|
| 1 | SVM_BP | 95,15 \pm 0,00 | SVM_BP | 96,86 \pm 0,00 | SVM_BP | 67,51 \pm 0,00 |
| 2 | SVM_C | 92,68 \pm 0,40 | SVM_C | 77,44 \pm 0,00 | SVM_C | 60,94 \pm 0,00 |
| 3 | SVM_CyH | 91,60 \pm 0,91 | C-SVM_V_A-Corr | 75,31 \pm 0,22 | SVM_CyH | 59,86 \pm 1,29 |
| 4 | C-SVM_V_A-Corr | 90,99 \pm 0,41 | SVM_CyH | 73,71 \pm 1,03 | C-SVM_V_A-Corr | 34,63 \pm 0,29 |

9.3. Reconocimiento de actividades físicas

En esta sección se realiza el análisis de las trazas de las matrices de confusión de las HMM con todas las configuraciones de sensores. En las tablas 13, 14 y 15 se muestran los resultados de cada modalidad de sensor teniendo en cuenta que la entrada del clasificador corresponde a la etiquetas entregadas por las SVM con el tamaño de ventana definido por *Andrés et al* [13]. En este caso, el desempeño del clasificador con el Kinect One® o con la red de sensores IMU en la mayoría de las clases es superior al 90 % exceptuando el caso de las clases 2 y 3 con el Kinect One® de los métodos HMM_CyH y HMM_C y las clases 1,9 y 10 de HMM_BP. A su vez, con la red de sensores IMU las clases 3 y 5 de HMM_CyH y HMM_C junto con las clases 9 y 10 de HMM_BP tienen eficiencias menores al 80 %. Por otro lado, con la red de sensores EMG aunque se tiene un desempeño aceptable, se observa que la clase 7 con los tres métodos presenta el mayor porcentaje de acierto en comparación con las otras clases superior al 83 %. Esto demuestra que utilizando una sola modalidad de sensor es posible realizar el reconocimiento de las actividades físicas, sin embargo, considerando los resultados en el análisis de los movimientos primitivos la probabilidad de equivocarse en la

estimación de las etiquetas para cada actividad es mínima para el caso del Kinect One® y los IMUs contrario al caso de los EMGs.

Tabla 13. Matriz de confusión para el reconocimiento de actividades físicas con el Kinect One®.

| Clases | Kinect One®-HMM_CyH | Kinect One®-HMM_C | Kinect One®-HMM_BP |
|----------|---------------------|-------------------|--------------------|
| 1 | 100.00 \pm 0.00 | 100.00 \pm 0.00 | 84.04 \pm 6.66 |
| 2 | 67.47 \pm 11.47 | 54.43 \pm 23.56 | 97.94 \pm 5.31 |
| 3 | 83.48 \pm 2.23 | 56.50 \pm 29.06 | 99.42 \pm 1.71 |
| 4 | 95.65 \pm 4.35 | 100.00 \pm 0.00 | 98.25 \pm 3.71 |
| 5 | 100.00 \pm 0.00 | 91.76 \pm 4.87 | 98.92 \pm 3.51 |
| 6 | 100.00 \pm 0.00 | 100.00 \pm 0.00 | 99.03 \pm 3.02 |
| 7 | 97.22 \pm 2.78 | 100.00 \pm 0.00 | 99.32 \pm 2.03 |
| 8 | 100.00 \pm 0.00 | 100.00 \pm 0.00 | 96.96 \pm 7.99 |
| 9 | 100.00 \pm 0.00 | 100.00 \pm 0.00 | 20.46 \pm 21.76 |
| 10 | 100.00 \pm 0.00 | 100.00 \pm 0.00 | 48.24 \pm 13.24 |
| Promedio | 94.38 \pm 0.30 | 90.27 \pm 3.51 | 84.26 \pm 2.99 |

Tabla 14. Matriz de confusión para el reconocimiento de actividades físicas con IMUs.

| Clases | IMUs-HMM_CyH | IMUs-HMM_C | IMUs-HMM_BP |
|----------|-------------------|-------------------|-------------------|
| 1 | 100.00 \pm 0.00 | 100.00 \pm 0.00 | 97.82 \pm 2.53 |
| 2 | 97.21 \pm 2.79 | 74.43 \pm 18.41 | 100.00 \pm 0.00 |
| 3 | 55.25 \pm 15.20 | 70.65 \pm 23.72 | 100.00 \pm 0.00 |
| 4 | 100.00 \pm 0.00 | 98.78 \pm 2.72 | 100.00 \pm 0.00 |
| 5 | 85.49 \pm 19.02 | 80.71 \pm 16.47 | 100.00 \pm 0.00 |
| 6 | 98.33 \pm 2.89 | 98.73 \pm 2.30 | 100.00 \pm 0.00 |
| 7 | 100.00 \pm 0.00 | 99.38 \pm 2.48 | 100.00 \pm 0.00 |
| 8 | 98.44 \pm 2.71 | 94.37 \pm 10.48 | 100.00 \pm 0.00 |
| 9 | 100.00 \pm 0.00 | 84.29 \pm 27.52 | 80.88 \pm 33.34 |
| 10 | 98.44 \pm 2.71 | 100.00 \pm 0.00 | 69.57 \pm 14.33 |
| Promedio | 93.32 \pm 0.71 | 90.13 \pm 4.08 | 94.83 \pm 4.29 |

Tabla 15. Matriz de confusión para el reconocimiento de actividades físicas con EMG.

| Clases | EMGs-HMM_CyH | EMGs-HMM_C | EMGs-HMM_BP |
|----------|-------------------|-------------------|-------------------|
| 1 | 56.00 \pm 11.41 | 44.84 \pm 10.87 | 83.67 \pm 12.22 |
| 2 | 66.81 \pm 12.91 | 85.09 \pm 12.27 | 74.20 \pm 16.10 |
| 3 | 39.36 \pm 11.46 | 44.84 \pm 11.09 | 59.99 \pm 12.32 |
| 4 | 82.63 \pm 9.55 | 87.60 \pm 9.47 | 83.99 \pm 10.87 |
| 5 | 75.63 \pm 13.03 | 76.66 \pm 10.98 | 93.70 \pm 6.59 |
| 6 | 72.43 \pm 8.13 | 84.89 \pm 9.15 | 84.62 \pm 8.75 |
| 7 | 83.14 \pm 10.17 | 90.67 \pm 7.22 | 91.03 \pm 6.31 |
| 8 | 76.97 \pm 8.21 | 83.62 \pm 10.35 | 90.54 \pm 8.40 |
| 9 | 61.20 \pm 27.80 | 52.98 \pm 22.50 | 26.66 \pm 17.49 |
| 10 | 59.99 \pm 8.56 | 59.80 \pm 15.80 | 40.69 \pm 23.91 |
| Promedio | 67.42 \pm 4.39 | 71.10 \pm 3.17 | 72.91 \pm 4.38 |

9.3.1. Experimento Kinect One®+IMUs

En la tabla 16 se muestran los resultados de la fusión de dos modalidades de sensores (Kinect One® y la red de sensores IMU). La clase 1 presenta un aumento del desempeño para todos los clasificadores del $100,00\% \pm 0,00\%$, mientras que con los modelos de HMM_CyH y HMM_C se tiene la misma eficiencia en las clases 4, 6 y 7 y con el modelo HMM_BP en la clase 3. Esto se debe a que la clasificación individual del Kinect One® o de los IMUs en esas clases presenta un alto desempeño.

Tabla 16. Matriz de confusión Kinect One®+IMUs.

| Clases | HMM_CyH | HMM_C | HMM_BP |
|----------|-------------------|-------------------|-------------------|
| 1 | 100.00 ± 0.00 | 100.00 ± 0.00 | 100.00 ± 0.00 |
| 2 | 73.89 ± 18.26 | 76.37 ± 20.35 | 96.53 ± 10.32 |
| 3 | 84.56 ± 17.11 | 84.68 ± 12.82 | 100.00 ± 0.00 |
| 4 | 100.00 ± 0.00 | 100.00 ± 0.00 | 85.84 ± 20.07 |
| 5 | 94.07 ± 4.61 | 93.42 ± 6.68 | 91.11 ± 19.89 |
| 6 | 100.00 ± 0.00 | 100.00 ± 0.00 | 95.87 ± 12.58 |
| 7 | 100.00 ± 0.00 | 100.00 ± 0.00 | 95.48 ± 15.00 |
| 18 | 99.33 ± 1.69 | 97.49 ± 6.87 | 94.23 ± 16.95 |
| 9 | 90.84 ± 23.16 | 83.68 ± 22.36 | 21.12 ± 18.94 |
| 10 | 81.03 ± 29.78 | 96.71 ± 13.96 | 11.40 ± 18.01 |
| Promedio | 92.37 ± 4.63 | 93.23 ± 3.73 | 79.16 ± 4.21 |

9.3.2. Experimento Kinect One®+EMGs

Los resultados de la fusión del Kinect One® con la red de sensores EMG se evidencian en la tabla 17. La clase 10 con el modelo de HMM_CyH tienen una eficiencia del $100,00\% \pm 0,00\%$ siendo esta actividad la de mejor desempeño con respecto a las demás clases y los otros modelos de clasificación propuestos. Sin embargo, las clases 2, 3, 4, 6 y 7 tienen un mejor desempeño con los modelos HMM_C y HMM_BP con respecto al modelo HMM_CyH. Esto se debe al desempeño que tuvieron ambos métodos en la clasificación de los movimientos primitivos.

Tabla 17. Matriz de confusión Kinect One®+EMGs.

| Clases | HMM_CyH | HMM_C | HMM_BP |
|----------|-------------------|-------------------|-------------------|
| 1 | 68.30 ± 16.78 | 45.09 ± 14.74 | 88.17 ± 15.97 |
| 2 | 74.35 ± 11.61 | 87.84 ± 13.44 | 79.44 ± 16.90 |
| 3 | 48.69 ± 13.52 | 61.76 ± 16.68 | 72.49 ± 20.96 |
| 4 | 82.12 ± 22.41 | 85.68 ± 10.52 | 82.97 ± 16.03 |
| 5 | 92.18 ± 12.77 | 86.48 ± 13.46 | 97.94 ± 4.43 |
| 6 | 78.25 ± 15.85 | 83.37 ± 10.79 | 86.73 ± 12.33 |
| 7 | 89.14 ± 8.92 | 95.36 ± 6.05 | 91.95 ± 8.04 |
| 18 | 99.01 ± 2.80 | 93.26 ± 8.27 | 89.71 ± 15.85 |
| 9 | 85.92 ± 11.12 | 77.08 ± 26.78 | 33.19 ± 18.37 |
| 10 | 100.00 ± 0.00 | 96.88 ± 5.59 | 31.83 ± 19.81 |
| Promedio | 81.80 ± 4.34 | 81.28 ± 4.23 | 75.44 ± 6.37 |

9.3.3. Experimento IMUs+EMGs

En la tabla 18 se muestra que la fusión de estas dos modalidades de sensores mejora los desempeños de los tres modelos de aprendizaje en comparación con los resultados individuales observados en las tablas 14 y 15. Las clases 5 y 8 presentan un rendimiento superior al 92 % con los métodos HMM_BP y HMM_CyH.

Tabla 18. Matriz de confusión IMUs+EMGs.

| Clases | HMM_CyH | HMM_C | HMM_BP |
|----------|-------------------|-------------------|-------------------|
| 1 | $74,50 \pm 15,72$ | $58,43 \pm 15,48$ | $88,05 \pm 13,31$ |
| 2 | $81,25 \pm 11,70$ | $95,62 \pm 6,18$ | $85,38 \pm 13,10$ |
| 3 | $56,48 \pm 12,69$ | $73,53 \pm 12,86$ | $73,71 \pm 20,89$ |
| 4 | $90,02 \pm 7,62$ | $89,51 \pm 9,34$ | $88,84 \pm 9,43$ |
| 5 | $93,99 \pm 7,18$ | $94,25 \pm 6,37$ | $96,78 \pm 5,89$ |
| 6 | $86,25 \pm 9,84$ | $88,58 \pm 9,50$ | $86,12 \pm 15,95$ |
| 7 | $87,69 \pm 10,80$ | $92,49 \pm 9,81$ | $95,11 \pm 5,40$ |
| 8 | $98,10 \pm 3,52$ | $92,19 \pm 8,91$ | $97,55 \pm 3,16$ |
| 9 | $88,73 \pm 11,17$ | $82,13 \pm 24,37$ | $66,65 \pm 39,74$ |
| 10 | $99,80 \pm 1,06$ | $98,67 \pm 4,37$ | $70,47 \pm 19,22$ |
| Promedio | $85,68 \pm 2,84$ | $86,54 \pm 4,15$ | $84,87 \pm 5,60$ |

9.3.4. Experimento Kinect One®+IMUs+EMGs

De acuerdo al resultado individual de cada una de las clases al combinar las tres modalidades de sensores se evidencia que la mayoría de las clases mejoraron su eficiencia en comparación con los experimentos visualizados en las tablas 13, 14, 15, 16, 17 y 18. Sin embargo, las clases 9 y 10 tienen una bajo porcentaje de acierto con el modelo HMM_BP debido a la baja eficiencia que tiene el movimiento primitivo presente en las dos actividades cuando se generan las etiquetas con las SVM para el caso del Kinect One®. Por otro lado, las clases que mejor desempeño tienen son las 5, 7 y 8 con los tres modelos de clasificación con un porcentaje de acierto superior al 92 %. Los métodos HMM_CyH y HMM_C presentan un desempeño superior al 87 % en comparación con HMM_BP como se visualiza en la tabla 19.

Tabla 19. Matriz de confusión Kinect One®+IMUs+EMGs.

| Clases | HMM_CyH | HMM_C | HMM_BP |
|----------|-------------------|-------------------|-------------------|
| 1 | 74.56 ± 14.92 | 43.18 ± 11.80 | 91.33 ± 13.70 |
| 2 | 79.01 ± 12.74 | 92.15 ± 12.25 | 86.54 ± 13.88 |
| 3 | 65.63 ± 9.48 | 84.62 ± 11.05 | 78.85 ± 17.56 |
| 4 | 92.74 ± 6.72 | 91.58 ± 7.91 | 78.47 ± 19.52 |
| 5 | 96.95 ± 4.48 | 92.97 ± 7.45 | 97.77 ± 3.53 |
| 6 | 86.86 ± 7.89 | 86.50 ± 9.45 | 91.49 ± 10.19 |
| 7 | 93.08 ± 5.82 | 95.40 ± 4.49 | 93.02 ± 5.64 |
| 18 | 98.79 ± 2.82 | 93.06 ± 7.76 | 96.82 ± 5.79 |
| 9 | 84.97 ± 19.39 | 93.52 ± 9.68 | 16.21 ± 14.73 |
| 10 | 98.54 ± 3.55 | 99.42 ± 2.48 | 9.69 ± 13.08 |
| Promedio | 87.11 ± 3.48 | 87.24 ± 3.68 | 74.02 ± 4.02 |

9.3.5. Comparación del desempeño del clasificador con todo el conjunto de sensores para cada modelo de la sección 9.2

La tabla 20 muestra el valor promedio de la diagonal principal de las matrices de confusión de la metodología propuesta por *Andrés et al* [13]. La clase 8 tiene un desempeño superior al 98 % para casi todas las configuraciones de sensores con la excepción de los EMGs. Por otro lado, la fusión de la red de sensores EMGs con cualquiera de las otras dos modalidades de sensores Kinect One® o IMUs permite mejorar el rendimiento en la clasificación de las

actividades.

Tabla 20. Comparación del desempeño con diferentes sensores y combinaciones de estos con el modelo HMM_CyH.

| Sensores | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 |
|-----------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Kinect One® | 100,00 ± 0,00 | 67,47 ± 11,47 | 83,48 ± 2,23 | 95,65 ± 4,35 | 100,00 ± 0,00 | 100,00 ± 0,00 | 97,22 ± 2,78 | 100,00 ± 0,00 | 100,00 ± 0,00 | 100,00 ± 0,00 |
| IMUs | 100,00 ± 0,00 | 97,21 ± 2,79 | 55,25 ± 15,20 | 100,00 ± 0,00 | 85,49 ± 19,02 | 98,33 ± 2,89 | 100,00 ± 0,00 | 98,44 ± 2,71 | 100,00 ± 0,00 | 98,44 ± 2,71 |
| EMGs | 56,00 ± 11,41 | 66,81 ± 12,91 | 39,36 ± 11,46 | 82,63 ± 9,55 | 75,63 ± 13,03 | 72,43 ± 8,13 | 83,14 ± 10,17 | 76,97 ± 8,21 | 61,20 ± 27,80 | 59,99 ± 8,56 |
| Kinect One®+IMUs | 100,00 ± 0,00 | 73,89 ± 18,26 | 84,56 ± 17,11 | 100,00 ± 0,00 | 94,07 ± 4,61 | 100,00 ± 0,00 | 100,00 ± 0,00 | 99,33 ± 1,69 | 90,84 ± 23,16 | 81,03 ± 29,78 |
| Kinect One®+EMGs | 68,30 ± 16,78 | 74,35 ± 11,61 | 48,69 ± 13,52 | 82,12 ± 22,41 | 92,18 ± 12,77 | 78,25 ± 15,85 | 89,14 ± 8,92 | 99,01 ± 2,80 | 85,92 ± 11,12 | 100,00 ± 0,00 |
| IMUs+EMGs | 74,50 ± 15,72 | 81,25 ± 11,70 | 56,48 ± 12,69 | 90,02 ± 7,62 | 93,99 ± 7,18 | 86,25 ± 9,84 | 87,69 ± 10,80 | 98,10 ± 3,52 | 88,73 ± 11,17 | 99,80 ± 1,06 |
| Kinect One®+IMUs+EMGs | 74,56 ± 14,92 | 79,01 ± 12,74 | 65,63 ± 9,48 | 92,74 ± 6,72 | 96,95 ± 4,48 | 86,86 ± 7,89 | 93,08 ± 5,82 | 98,79 ± 2,82 | 84,97 ± 19,39 | 98,54 ± 3,55 |

En la tabla 21 se indican los valores promedio de la diagonal principal de las matrices de confusión de la metodología propuesta utilizando el método 1 de penalización (C-SVM) para la generación de etiquetas los movimientos primitivos. La clase 7 tiene un desempeño superior al 90 % para todas las configuraciones de sensores. Por otro lado, la fusión del Kinect One® con la red de sensores IMU mejora el rendimiento del clasificador aproximadamente en un 3 % en comparación con los resultados de cada sensor. Como se mencionó anteriormente, se observa de nuevo el incremento en el desempeño del clasificador de los EMGs con la fusión de las otras modalidades de sensor.

Tabla 21. Comparación del desempeño con diferentes sensores y combinaciones de estos con el modelo HMM_C.

| Sensores | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 |
|-----------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Kinect One® | 100,0 ± 0,00 | 54,43 ± 23,56 | 56,50 ± 29,06 | 100,00 ± 0,00 | 91,76 ± 4,87 | 100,00 ± 0,00 | 100,00 ± 0,00 | 100,00 ± 0,00 | 100,00 ± 0,00 | 100,00 ± 0,00 |
| IMUs | 100,00 ± 0,00 | 74,43 ± 18,41 | 70,65 ± 23,72 | 98,78 ± 2,72 | 80,71 ± 16,47 | 98,73 ± 2,30 | 99,38 ± 2,48 | 94,37 ± 10,48 | 84,29 ± 27,52 | 100,00 ± 0,00 |
| EMGs | 44,84 ± 10,87 | 85,09 ± 12,27 | 44,84 ± 11,09 | 87,60 ± 9,47 | 76,66 ± 10,98 | 84,89 ± 9,15 | 90,67 ± 7,22 | 83,62 ± 10,35 | 52,98 ± 22,50 | 59,80 ± 15,80 |
| Kinect One®+IMUs | 100,00 ± 0,00 | 76,37 ± 20,35 | 84,68 ± 12,82 | 100,00 ± 0,00 | 93,42 ± 6,68 | 100,00 ± 0,00 | 100,00 ± 0,00 | 97,49 ± 6,87 | 83,68 ± 22,36 | 96,71 ± 13,96 |
| Kinect One®+EMGs | 45,09 ± 14,74 | 87,84 ± 13,44 | 61,76 ± 16,68 | 85,68 ± 10,52 | 86,48 ± 13,46 | 83,37 ± 10,79 | 95,36 ± 6,05 | 93,26 ± 8,27 | 77,08 ± 26,78 | 96,88 ± 5,59 |
| IMUs+EMGs | 58,43 ± 15,48 | 95,62 ± 6,18 | 73,53 ± 12,86 | 89,51 ± 9,34 | 94,25 ± 6,37 | 88,58 ± 9,50 | 92,49 ± 9,81 | 92,19 ± 8,91 | 82,13 ± 24,37 | 98,67 ± 4,37 |
| Kinect One®+IMUs+EMGs | 43,18 ± 11,80 | 92,15 ± 12,25 | 84,62 ± 11,05 | 91,58 ± 7,91 | 92,97 ± 7,45 | 86,50 ± 9,45 | 95,40 ± 4,49 | 93,06 ± 7,76 | 93,52 ± 9,68 | 99,42 ± 2,48 |

En la tabla 22 se visualiza el valor promedio de la diagonal principal de las matrices de confusión de la metodología propuesta utilizando el método 2 de penalización (SVM binaria ponderada) para la generación de etiquetas los movimientos primitivos. La clase 5 tiene un desempeño superior al 91 % para todas las configuraciones de sensores. La fusión de las redes de sensores IMU y EMG mejora el rendimiento del clasificador aproximadamente en un 12 % en comparación con el promedio de los sensores EMGs para todas la clases. Por último, las clases 9 y 10 evidencian un desempeño menor al 50 % en las configuraciones Kinect One®, EMG, Kinect One®+IMUs y Kinect One®+IMUs+EMGs. Esto se debe al bajo desempeño

del clasificador en la generación de etiquetas con el Kinect One® para la identificación de los movimientos primitivos de la clase 16 la cual está presente en la ejecución de las actividades 9 y 10.

Tabla 22. Comparación del desempeño con diferentes sensores y combinaciones de estos con el modelo HMM_BP.

| Sensores | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 |
|-----------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Kinect One® | 84,04 ± 6,66 | 97,94 ± 5,31 | 99,42 ± 1,71 | 98,25 ± 3,71 | 98,92 ± 3,51 | 99,03 ± 3,02 | 99,32 ± 2,03 | 96,96 ± 7,99 | 20,462 ± 1,76 | 48,24 ± 13,24 |
| IMUs | 97,82 ± 2,53 | 100,00 ± 0,00 | 100,00 ± 0,00 | 100,00 ± 0,00 | 100,00 ± 0,00 | 100,00 ± 0,00 | 100,00 ± 0,00 | 100,00 ± 0,00 | 80,88 ± 33,34 | 69,57 ± 14,33 |
| EMGs | 83,67 ± 12,22 | 74,20 ± 16,10 | 59,99 ± 12,32 | 83,99 ± 10,87 | 93,70 ± 6,59 | 84,62 ± 8,75 | 91,03 ± 6,31 | 90,54 ± 8,40 | 26,66 ± 17,49 | 40,69 ± 23,91 |
| Kinect One®+IMUs | 100,00 ± 0,00 | 96,53 ± 10,32 | 100,00 ± 0,00 | 85,84 ± 20,07 | 91,11 ± 19,89 | 95,87 ± 12,58 | 95,481 ± 5,00 | 94,23 ± 16,95 | 21,12 ± 18,94 | 11,40 ± 18,01 |
| Kinect One®+EMGs | 88,17 ± 15,97 | 79,44 ± 16,90 | 72,49 ± 20,96 | 82,97 ± 16,03 | 97,94 ± 4,43 | 86,73 ± 12,33 | 91,95 ± 8,04 | 89,71 ± 15,85 | 33,19 ± 18,37 | 31,83 ± 19,81 |
| IMUs+EMGs | 88,05 ± 13,31 | 85,38 ± 13,10 | 73,71 ± 20,89 | 88,84 ± 9,43 | 96,78 ± 5,89 | 86,12 ± 15,95 | 95,11 ± 5,40 | 97,55 ± 3,16 | 66,65 ± 39,74 | 70,47 ± 19,22 |
| Kinect One®+IMUs+EMGs | 91,33 ± 13,70 | 86,54 ± 13,88 | 78,85 ± 17,56 | 78,47 ± 19,52 | 97,77 ± 3,53 | 91,49 ± 10,19 | 93,02 ± 5,64 | 96,82 ± 5,79 | 16,21 ± 14,73 | 9,69 ± 13,08 |

De acuerdo a los resultados experimentales de las tablas 23 y 24 se observa que la metodología propuesta con el segundo método de penalización presenta la mayor eficiencia en el reconocimiento de las actividades físicas con un porcentaje de acierto del 94,83% utilizando la red de sensores IMU. Este valor es comparable con la clasificación de las actividades usando el Kinect One®, Kinect One®+IMUs, Kinect One®+EMGs, IMUs+EMGs y Kinect One®+IMUs+EMGs. Por otro lado, se evidencia que los dos métodos de penalización C-SVM y SVM binaria ponderada mejoran el desempeño del clasificador con los datos EMG. Sin embargo, este método en comparación con el enfoque de *Andrés et al* [13] presenta un rendimiento aceptable con la fusión de las diferentes modalidades de sensores.

Tabla 23. Comportamiento promedio de la metodología con una modalidad de sensor para cada modelo de clasificación.

| Puesto | Kinect One® | | IMUs | | EMGs | |
|--------|-------------|--------------|---------|--------------|---------|--------------|
| 1 | HMM_CyH | 94.38 ± 0.30 | HMM_BP | 94.83 ± 4.29 | HMM_BP | 72.91 ± 4.38 |
| 2 | HMM_C | 90.27 ± 3.51 | HMM_CyH | 93.32 ± 0.71 | HMM_C | 71.10 ± 3.17 |
| 3 | HMM_BP | 84.26 ± 2.99 | HMM_C | 90.13 ± 4.08 | HMM_CyH | 67.42 ± 4.39 |

Tabla 24. Comportamiento promedio de la metodología con la fusión de sensores para cada modelo de clasificación.

| Puesto | Kinect One®+IMUs | | Kinect One®+EMGs | | IMUs+EMGs | | Kinect One®+IMUs+EMGs | |
|--------|------------------|--------------|------------------|--------------|-----------|--------------|-----------------------|--------------|
| 1 | HMM_C | 93.23 ± 3.73 | HMM_CyH | 81.80 ± 4.34 | HMM_C | 86.54 ± 4.15 | HMM_C | 87.24 ± 3.68 |
| 2 | HMM_CyH | 92.37 ± 4.63 | HMM_C | 81.28 ± 4.23 | HMM_CyH | 85.68 ± 2.84 | HMM_CyH | 87.11 ± 3.48 |
| 3 | HMM_BP | 79.16 ± 4.21 | HMM_BP | 75.44 ± 6.37 | HMM_BP | 84.87 ± 5.60 | HMM_BP | 74.02 ± 4.02 |

10. CONCLUSIONES Y RECOMENDACIONES

10.1. Conclusiones

En este trabajo se realizó un estudio comparativo articulando diferentes modelos de aprendizaje para la identificación de los movimientos primitivos, donde se utiliza la metodología propuesta en [13] con el fin de plantear dos métodos que penalicen las máquinas de vectores de soporte para su respectiva comparación. En este caso, se comprobó que con cualquiera de los dos métodos de penalización planteados en este documento aumenta el desempeño de los clasificadores con las tres modalidades de sensores en comparación con [13], siendo el método de la SVM binaria pondera el de mayor eficiencia con un porcentaje de acierto del 95,15 % con el Kinect One® y del 96,86 % con la red de sensores IMU, esto se debe a que el método tiene en cuenta el desbalance que existe entre las diferentes clases asociadas a los movimientos primitivos, penalizando el parámetro de regularización C para cada clase y considerando que el método one vs one presenta siempre un mejor rendimiento con respecto al método de one vs all. Por otra parte, con base a las etiquetas generadas por las SVMs con la secuencia de los sub-movimientos de cada acción, se utilizó la HMM para el reconocimiento de las actividades físicas, mostrando una mayor eficiencia todos los experimentos que utilizan a red de sensores IMU, en este caso el mejor desempeño se tuvo también con los parámetros entregados a la HMM por el segundo método de penalización con los datos correspondientes a la red de sensores IMU, con un porcentaje de acierto del 94,83 % en su clasificación.

Se comprobó que la fusión de diferentes modalidades de sensores permite mejorar el desempeño del clasificador, específicamente esto se evidencia con la red de sensores EMG, los cuales tienen una baja eficiencia con cualquiera de los tres modelos comparativos (HMM_CyH, HMM_C y HMM_BP), y mejora aproximadamente en un 10 % con la siguientes combinaciones de sensores Kinect One®+EMGs IMUs+EMGs Kinect One®+IMUs+EMGs. Sin embargo, para el reconocimiento de actividades físicas la mayor tasa de desempeño no se obtuvo con la fusión de estas tres modalidades de sensores, esto se debe a que algunas clases se encuentran traslapadas de acuerdo al sensor utilizado, siendo los datos de los sensores EMG los más traslapados.

También se amplió la base de datos de actividades físicas desarrollada por [13] en 10 actividades con el registro sincronizado del Kinect One®, IMUs y EMGs, donde se incluyeron más puntos articulados del cuerpo humano y 16 movimientos primitivos, por lo que se puede ampliar la base de datos con actividades combinadas, teniendo en cuenta el conjunto de movimientos primitivos definidos en este documento, como por ejemplo levantarse de la silla y alzar la mano derecha, realizar el seguimiento de la marcha iniciado en una posición de reposo, para luego caminar y trotar, entre otras, esto con el objetivo de forzar el entrenamiento de la HMM en la identificación de secuencias más grandes de movimientos primitivos con varios estados. Además, es importante determinar de forma adecuada la cantidad y el tiempo de ejecución de cada uno de los sub-movimientos presentes en una actividad considerando la separabilidad que debe existir entre estos y su duración, ya que de esto depende encontrar un tamaño óptimo de ventana para el proceso de segmentación. En este trabajo se establecen dos tamaños de ventanas para la identificación de los movimientos primitivos, donde la primera se encuentra de forma experimental y la segunda mediante el análisis de la autocorrelación de las señales de cada uno de los sensores, obteniendo un resultado comparable del Kinect One® y los IMUs con la ventana experimental, pero presentado mejor eficiencia el primer caso.

10.2. Recomendaciones

En la construcción de la base de datos es conveniente realizar un estudio sobre la cantidad mínima de sensores (IMU y EMG) y la ubicación de estos en el cuerpo humano para obtener el registro completo de los diferentes movimientos asociados a las actividades, esto con el fin de determinar los sub-movimientos con los rangos de tiempo asociados a la ejecución de cada una de las actividades. Además se recomienda utilizar sensores de movimiento muscular como las Myo, el cual es un brazalete que integra sensores electromiográficos y acelerómetros con menores pérdidas de la señal o desconexión del sistema de medida en comparación con los electrodos.

Un factor importante para el reconocimiento de los movimientos primitivos es la etiquetación de su base de datos, en este caso se realizó de forma manual teniendo errores humanos,

para lo cual se propone utilizar métodos de clasificación no supervisados como para la generación automática de las etiquetas, ya que este proceso afecta directamente el desempeño del clasificador en la identificación de los sub-movimientos.

Siendo C el parámetro de regularización de la SVM que se utilizó para realizar la penalización, es importante explorar otras metodologías aparte de tener una grilla de observación en un rango determinado que estime un valor óptimo de este parámetro para cada una de las clases del modelo de entrenamiento, efectuando una búsqueda aleatoria entre los datos de observación o mediante un método de optimización como la bayesiana, para encontrar este valor.

BIBLIOGRAFÍA

- [1] CALVO, Andrés. RECONOCIMIENTO AUTOMÁTICO DE ACTIVIDADES FÍSICAS HUMANAS EN SISTEMAS MULTIMODALES. Tesis de Maestría, UNIVERSIDAD TECNOLÓGICA DE PEREIRA, 2015. (document), 1.1, 1.2, 2, 3.2, 3.8, 4.1, 5, 5.1, 5.1
- [2] LEWIS, JP. A short SVM (support vector machine) tutorial. En: CGIT Lab/IMSC, University Southern California, 2004. (document), 4.1
- [3] SUÁREZ, Enrique J Carmona. Tutorial sobre máquinas de vectores soporte (sVM). En: Tutorial sobre Máquinas de Vectores Soporte (SVM), 2014. (document), 4.1.1
- [4] GIL GONZÁLEZ, Walter Julián, *et al.* Clasificador robusto basado en máquinas de soporte vectorial para la localización de fallas en sistemas de distribución. Tesis de Maestría, Pereira: Universidad Tecnológica de Pereira, 2013. (document), 4.1.2
- [5] SCHOLKOPF, Bernhard y SMOLA, Alexander J. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press, Cambridge, MA, USA, 2001. ISBN 0262194759. (document), 4.1.3
- [6] SHAO, S., *et al.* Automatic EEG Artifact Removal: A Weighted Support Vector Machine Approach With Error Correction. En: IEEE Transactions on Biomedical Engineering, tomo 56, n^o 2, 2009, págs. 336–344. ISSN 0018-9294. (document), 4.1.4, 7.2
- [7] MELO LOZADA, Cristian Andrés; ORTEGA DEL VECCHIO, Luisa Fernanda, *et al.* HMM espectrales para el reconocimiento de emociones. B.S. thesis, Pereira: Universidad Tecnológica de Pereira, 2015. (document), 4.2
- [8] CAO, Liang, *et al.* GCHAR: An efficient Group-based Context aware human activity recognition on smartphone. En: Journal of Parallel and Distributed Computing, 2017. ISSN 0743-7315. 1, 1.1, 1.2, 5.1

- [9] KHAN, Aftab, *et al.* Optimising sampling rates for accelerometer-based human activity recognition. En: Pattern Recognition Letters, tomo 73, n^o Supplement C, 2016, págs. 33 – 40. ISSN 0167-8655. 1, 1.1, 1.2, 2

- [10] GRAVINA, Raffaele, *et al.* Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges. En: Information Fusion, tomo 35, n^o Supplement C, 2017, págs. 68 – 80. ISSN 1566-2535. 1, 1.1, 1.2, 2, 3.5

- [11] CHEN, Yen-Lun, *et al.* Dimensionality reduction of data sequences for human activity recognition. En: Neurocomputing, tomo 210, n^o Supplement C, 2016, págs. 294 – 302. ISSN 0925-2312. 1, 1.1, 1.2

- [12] TAKANO, Wataru; IMAGAWA, Hirotaka y NAKAMURA, Yoshihiko. Spatio-temporal structure of human motion primitives and its application to motion prediction. En: Robotics and Autonomous Systems, tomo 75, n^o Part B, 2016, págs. 288 – 296. ISSN 0921-8890. 1, 7

- [13] CALVO, Andres; HOLGUIN, German A y MEDEIROS, Henry. Human activity recognition using multi-modal data fusion with Kinect, IMUs and EMGs. En: The 23 Iberoamerican Congress on Pattern Recognition, CIARP 2018, tomo 27, n^o 11, 2018. 1, 1.1, 2, 5, 5.1, 6, 7, 1, 7.2, 9, 1, 2, 9.1, 1, 9.1, 9.2, 9.2, 9.2, 9.3, 9.3.5, 9.3.5, 10.1

- [14] MORALES, Santiago. IDENTIFICACIÓN DE ACTIVIDAD HUMANA USANDO APRENDIZAJE NO SUPERVISADO EN SISTEMAS MULTIMODALES. Tesis de Maestría, UNIVERSIDAD TECNOLÓGICA DE PEREIRA, 2016. 1, 1.1, 1.2, 2, 3.1, 7

- [15] WANG, B.; YANG, C. y XIE, Q. Human-machine interfaces based on EMG and Kinect applied to teleoperation of a mobile humanoid robot. En: Proceedings of the 10th World Congress on Intelligent Control and Automation, 2012, págs. 3903–3908. 1

- [16] FENG, S. y MURRAY-SMITH, R. Fusing Kinect sensor and inertial sensors with multi-rate Kalman filter. En: IET Conference on Data Fusion Target Tracking 2014: Algorithms and Applications (DF TT 2014), 2014, págs. 1–8. 1
- [17] XIANG, C., *et al.* Comparing Real-Time Human Motion Capture System Using Inertial Sensors with Microsoft Kinect. En: 2014 7th International Conference on Ubi-Media Computing and Workshops, 2014, págs. 53–58. 1
- [18] CAON, Maurizio, *et al.* Kinesiologic Electromyography for Activity Recognition. En: Proceedings of the 6th International Conference on Pervasive Technologies Related to Assistive Environments, PETRA '13. ACM, New York, NY, USA, 2013. ISBN 978-1-4503-1973-7, págs. 34:1–34:7. URL <http://doi.acm.org/10.1145/2504335.2504372>. 1, 5
- [19] KOSKIMAKI, Heli y SIIRTOLA, Pekka. Accelerometer vs. Electromyogram in Activity Recognition. En: ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal, tomo 5, n^o 3, 2016. ISSN 2255-2863. 1
- [20] VICENTE, Isabel Serrano, *et al.* Action recognition and understanding through motor primitives. En: Advanced Robotics, tomo 21, n^o 15, 2007, págs. 1687–1707. 1, 6, 7
- [21] LIN, J. F. S.; KARG, M. y KULIĆ, D. Movement Primitive Segmentation for Human Motion Modeling: A Framework for Analysis. En: IEEE Transactions on Human-Machine Systems, tomo 46, n^o 3, 2016, págs. 325–339. ISSN 2168-2291. 1, 3.1, 3.4, 3.6, 5.1, 7
- [22] RENGIFO, Daniel y GALLEGU, Jhon. COMPARACIÓN DE TÉCNICAS DE REDUCCIÓN DE DIMENSIONALIDAD PARA LA CLASIFICACIÓN DE ACTIVIDADES FÍSICAS HUMANAS UTILIZANDO MÉTODOS ESTADÍSTICOS. Inf. téc., UNIVERSIDAD TECNOLÓGICA DE PEREIRA, 2016. 1, 1.1, 3.7
- [23] GAGLIO, S.; RE, G. L. y MORANA, M. Human Activity Recognition Process Using 3-D Posture Data. En: IEEE Transactions on Human-Machine Systems, tomo 45, n^o 5, 2015, págs. 586–597. ISSN 2168-2291. 1, 1.1, 1.2, 2, 7

- [24] CASTELLANI, Andrea, *et al.* Hybrid HMM/SVM model for the analysis and segmentation of teleoperation tasks. En: Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on, tomo 3. IEEE, 2004, págs. 2918–2923. 1, 6
- [25] JIANG, Min, *et al.* Informative joints based human action recognition using skeleton contexts. En: Signal Processing: Image Communication, tomo 33, nº Supplement C, 2015, págs. 29 – 40. ISSN 0923-5965. 1.1, 1.2, 3.8
- [26] QIAO, Ruizhi, *et al.* Learning discriminative trajectorylet detector sets for accurate skeleton-based action recognition. En: Pattern Recognition, tomo 66, nº Supplement C, 2017, págs. 202 – 212. ISSN 0031-3203. 1.1, 1.2
- [27] AKBARI, A.; THOMAS, X. y JAFARI, R. Automatic noise estimation and context-enhanced data fusion of IMU and Kinect for human motion measurement. En: 2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN), 2017, págs. 178–182. 1.1, 1.2, 2
- [28] ZHANG, Mi y SAWCHUK, Alexander. Motion Primitive-Based Human Activity Recognition Using a Bag-of-Features Approach. En: Life and Medical Sciences, 2012. 1.1, 1.2, 2, 7
- [29] SADRI, Amin; REN, Yongli y SALIM, Flora D. Information gain-based metric for recognizing transitions in human activities. En: Pervasive and Mobile Computing, tomo 38, nº Part 1, 2017, págs. 92 – 109. ISSN 1574-1192. 1.1, 2
- [30] RONAÑO, Charissa Ann y CHO, Sung-Bae. Human activity recognition with smartphone sensors using deep learning neural networks. En: Expert Systems with Applications, tomo 59, nº Supplement C, 2016, págs. 235 – 244. ISSN 0957-4174. 2
- [31] CHENG, Guangchun, *et al.* Advances in Human Action Recognition: A Survey. En: CoRR, tomo abs/1501.05964, 2015. 2, 5
- [32] MICROSOFT. Kinect for Windows Technical Datasheet. Partner Network, 3^a ed^{ón}. 3.3

- [33] LEWIS, J.P. A Short SVM (Support Vector Machine) Tutorial. Inf. téc., CGIT Lab / IMSC, 2004. URL <http://www.idiom.com/~zilla/Work/Notes/svmtutorial.pdf>. 3.9.1, 3.9.2
- [34] VALENCIA, Tomas Echeverri y GONZÁLEZ, Julián Gil. Comparación de técnicas de clasificación de múltiples anotadores para la valoración Automática de la calidad de voz. Inf. téc., Universidad Tecnológica de Pereira, 2014. 3.10
- [35] MORANA, Marco. KARD - Kinect Activity Recognition Dataset, 2017. URL <https://data.mendeley.com/datasets/k28dtm7tr6/1>. 5
- [36] ROGGEN, Daniel, *et al.* Collecting complex activity datasets in highly rich networked sensor environments. En: Networked Sensing Systems (INSS), 2010 Seventh International Conference on. IEEE, 2010, págs. 233–240. 5
- [37] BANOS, Oresti, *et al.* Dealing with the effects of sensor displacement in wearable activity recognition. En: Sensors, tomo 14, nº 6, 2014, págs. 9995–10023. 5
- [38] SHAHROUDY, Amir, *et al.* NTU RGB+D: A Large Scale Dataset for 3D Human Activity Analysis. En: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016. 5, 5.1
- [39] OFLI, Ferda, *et al.* Berkeley mhad: A comprehensive multimodal human action database. En: Applications of Computer Vision (WACV), 2013 IEEE Workshop on. IEEE, 2013, págs. 53–60. 5, 5.1
- [40] IENENG CHEN. 3D Human Recognition Dataset, 2018. URL <https://beckschen.github.io/2018/05/13/HA-Dataset/>. 5
- [41] ANGUITA, Davide, *et al.* A public domain dataset for human activity recognition using smartphones. En: ESANN, 2013. 5
- [42] KRAFT, Marek, *et al.* Toward evaluation of visual navigation algorithms on RGB-D data from the first- and second-generation Kinect. En: Machine Vision and Applications, tomo 28, nº 1, 2017, págs. 61–74. ISSN 1432-1769. 5.1

- [43] WASENMÜLLER, Oliver y STRICKER, Didier. Comparison of Kinect V1 and V2 Depth Images in Terms of Accuracy and Precision. En: ACCV Workshops, 2016. 5.1
- [44] ZENNARO, S., *et al.* Performance evaluation of the 1st and 2nd generation Kinect for multimedia applications. En: 2015 IEEE International Conference on Multimedia and Expo (ICME), 2015, págs. 1–6. 5.1, 7.1.2
- [45] SARBOLANDI, Hamed; LEFLOCH, Damien y KOLB, Andreas. Kinect Range Sensing: Structured-Light versus Time-of-Flight Kinect. En: Computer Vision and Image Understanding, tomo 139, 2015, págs. 1–20. 5.1
- [46] PHAM, Huy-Hieu, *et al.* Exploiting deep residual networks for human action recognition from skeletal data. En: Computer Vision and Image Understanding, 2018. 5.1
- [47] ROLDÁN JIMÉNEZ, Cristina, *et al.* Estudio de la cinemática del miembro superior e inferior mediante sensores inerciales. Tesis Doctoral, Universidad de Málaga, 2017. 5.2
- [48] WINTER, David A. Moments of force and mechanical power in jogging. En: Journal of biomechanics, tomo 16, nº 1, 1983, págs. 91–97. 5.2
- [49] DUGAN, Sheila A y BHAT, Krishna P. Biomechanics and analysis of running gait. En: Physical Medicine and Rehabilitation Clinics, tomo 16, nº 3, 2005, págs. 603–621. 5.2
- [50] BERNAL, Antonio Gomez; BECERRO-DE BENGEOA-VALLEJO, Ricardo y LOSA-IGLESIAS, Marta Elena. Reliability of the OptoGait portable photoelectric cell system for the quantification of spatial-temporal parameters of gait in young adults. En: Gait & posture, tomo 50, 2016, págs. 196–200. 5.2
- [51] NAKAGAWA, J., *et al.* Analysis of Joint Correlation between Arm and Lower Body in Dart Throwing Motion. En: 2013 IEEE International Conference on Systems, Man, and Cybernetics, 2013. ISSN 1062-922X, págs. 1223–1228. 5.2
- [52] HUANG, Xuedong, *et al.* Spoken language processing: A guide to theory, algorithm, and system development, tomo 1. Prentice hall PTR Upper Saddle River, 2001. 6

- [53] HANNAFORD, Blake y LEE, Paul. Hidden Markov model analysis of force/torque information in telemanipulation. En: The International journal of robotics research, tomo 10, nº 5, 1991, págs. 528–539. 6
- [54] KULIC, Dana; TAKANO, Wataru y NAKAMURA, Yoshihiko. Online segmentation and clustering from continuous observation of whole body motions. En: IEEE Transactions on Robotics, tomo 25, nº 5, 2009, pág. 1158. 7
- [55] PAGLIARI, Diana y PINTO, Livio. Calibration of Kinect for Xbox One and Comparison between the Two Generations of Microsoft Sensors. En: Sensors, tomo 15, nº 11, 2015, págs. 27569–27589. ISSN 1424-8220. 7.1.2
- [56] WU, Haitao, *et al.* Human activity recognition based on the combined svm&hmm. En: Information and Automation (ICIA), 2014 IEEE International Conference on. IEEE, 2014, págs. 219–224. 7.1.2
- [57] ZHANG, Mi y SAWCHUK, Alexander A. A feature selection-based framework for human activity recognition using wearable multimodal sensors. En: Proceedings of the 6th International Conference on Body Area Networks. ICST, 2011, págs. 92–98. 7.1.3
- [58] RINCÓN, Jhon Fredy Gallego y ALMANZA, Daniel Fernando Rengifo. Comparación de técnicas de reducción de dimensionalidad para la clasificación de actividades físicas humanas utilizando métodos estadísticos. Universidad Tecnológica de Pereira. Facultad de Ingenierías Eléctrica, 2016. 7.2
- [59] BECKER, Natalia, *et al.* penalizedSVM: a R-package for Feature Selection SVM Classification. En: Bioinformatics, tomo 25, 2009, págs. 1711–1712. 7.2

11. ANEXOS

11.1. Matrices de Confusión de los movimientos primitivos

11.1.1. Resultados experimentales con la metodología de SVM_CyH

Figura 15. Matriz de confusión Movimientos primitivos con Kinect One®.

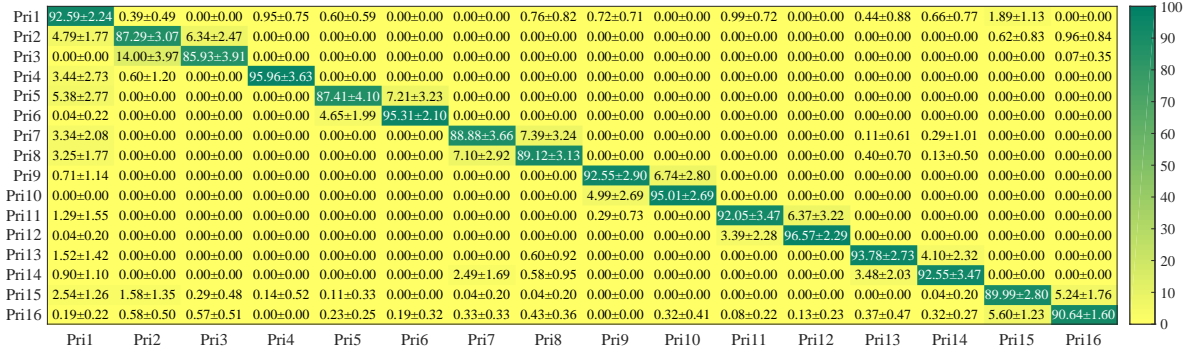


Figura 16. Matriz de confusión Movimientos primitivos con IMUs.

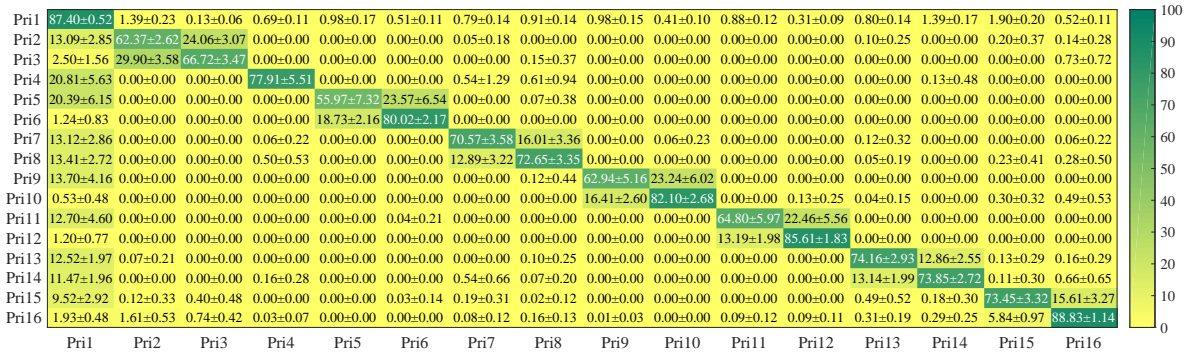
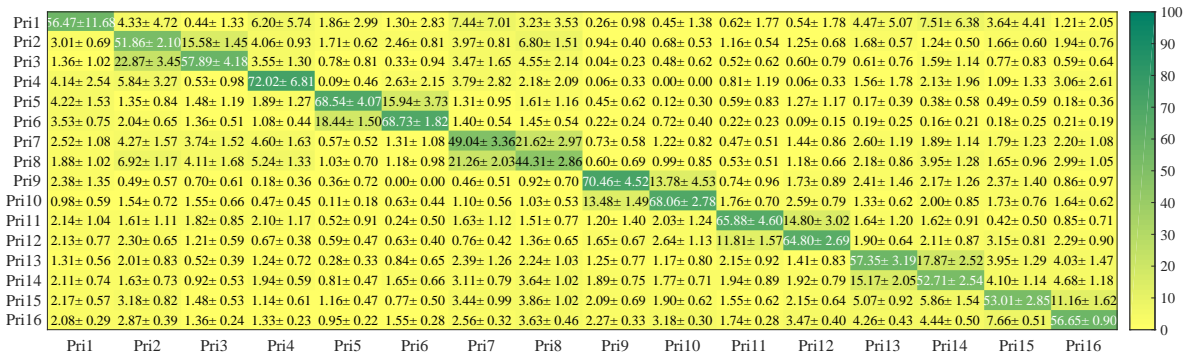


Figura 17. Matriz de confusión Movimientos primitivos con EMGs.



11.1.2. Resultados experimentales con la metodología de SVM_C

Figura 18. Matriz de confusión Movimientos primitivos con Kinect One®.

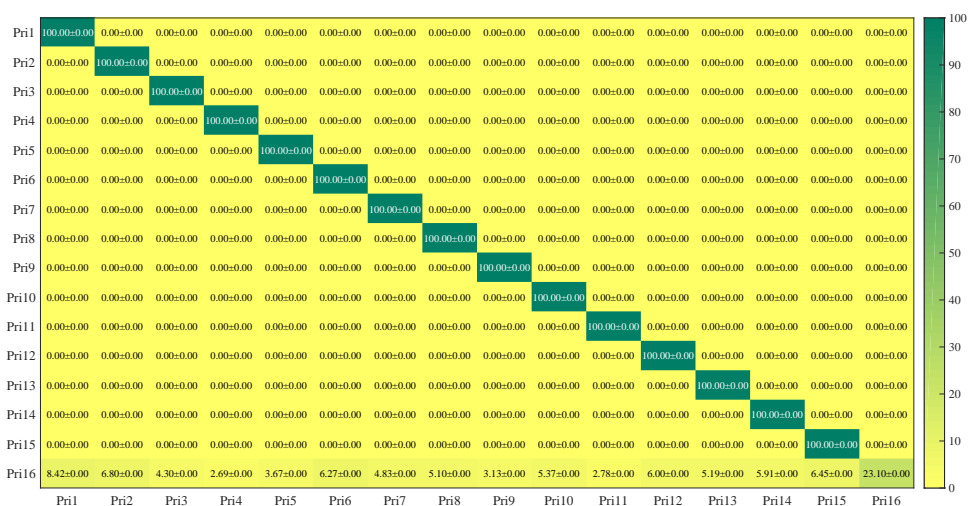


Figura 19. Matriz de confusión Movimientos primitivos con IMUs.

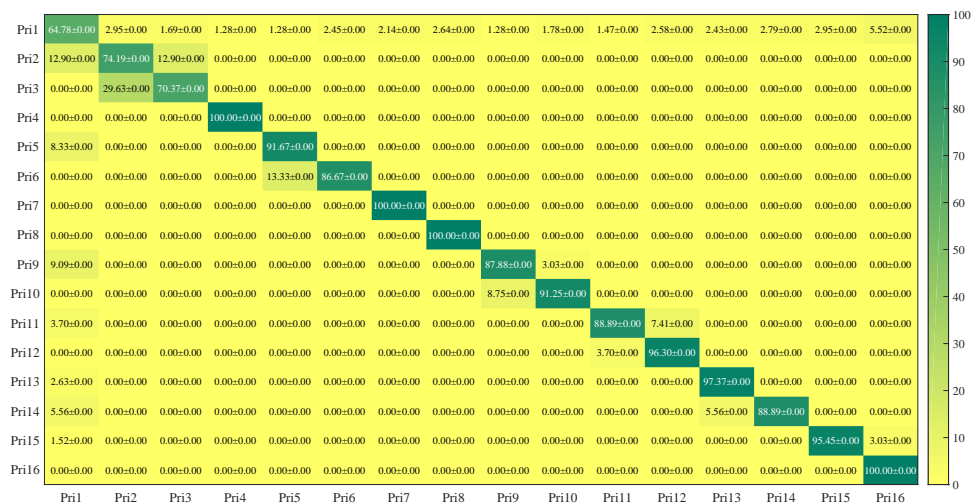
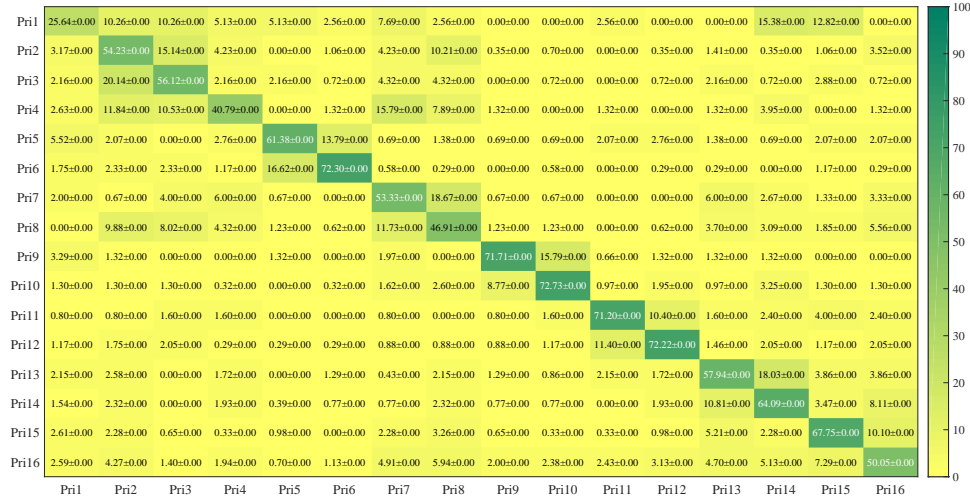


Figura 20. Matriz de confusión Movimientos primitivos con EMGs.



11.1.3. Resultados experimentales con la metodología de SVM_BP

Figura 21. Matriz de confusión Movimientos primitivos con Kinect One®.

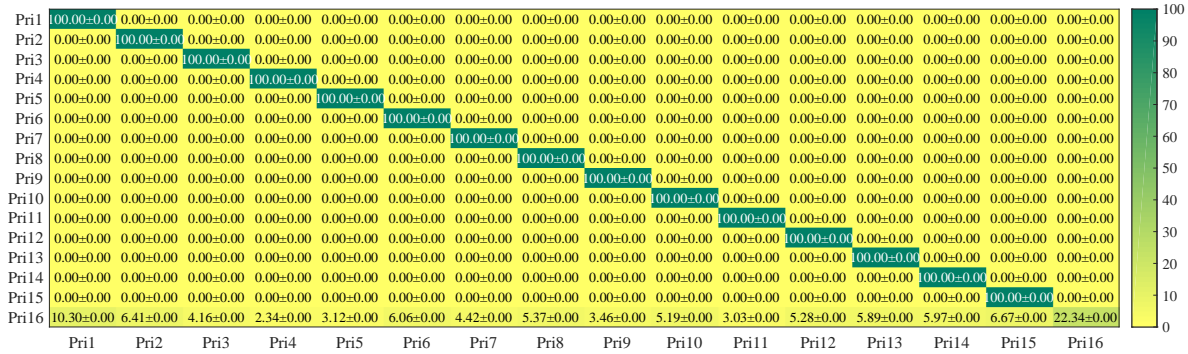


Figura 22. Matriz de confusión Movimientos primitivos con IMUs.

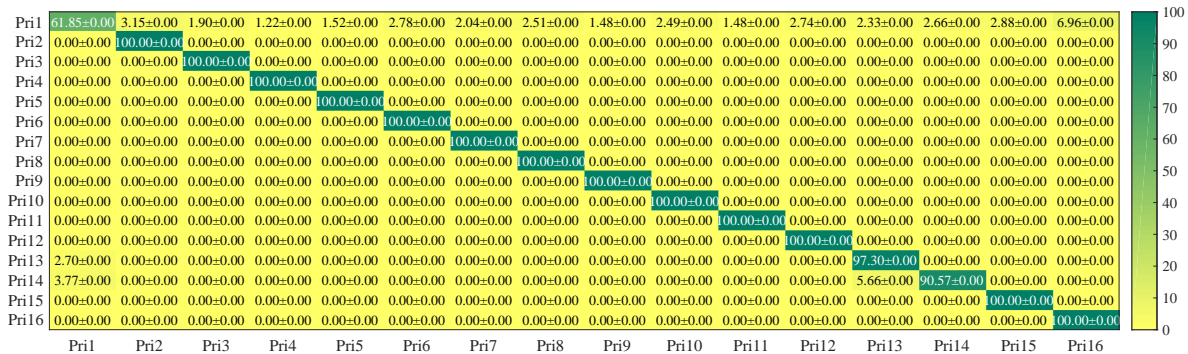
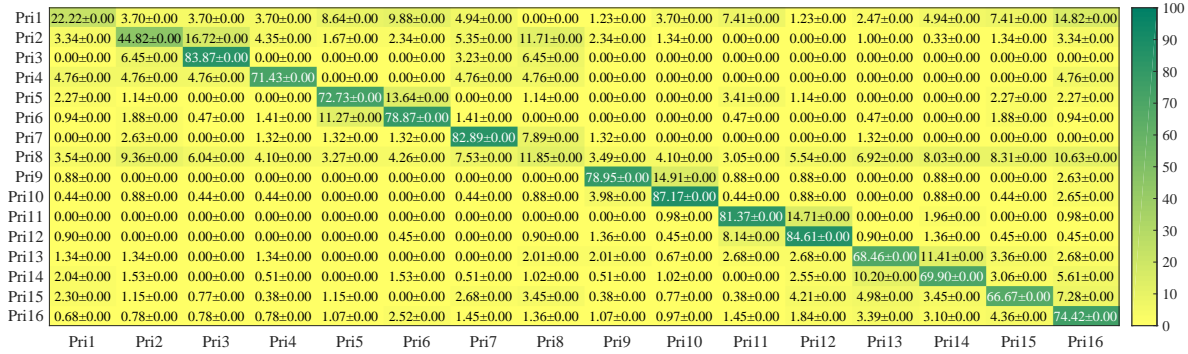


Figura 23. Matriz de confusión Movimientos primitivos con EMGs.



11.1.4. Resultados experimentales con la metodología de C-SVM_V_A-Corr

Figura 24. Matriz de confusión Movimientos primitivos con Kinect One®.

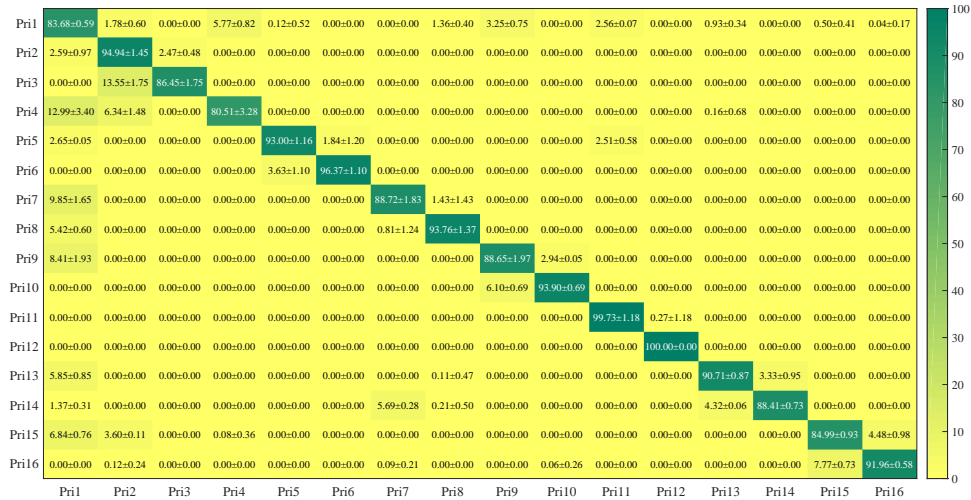


Figura 25. Matriz de confusión Movimientos primitivos con IMUs.

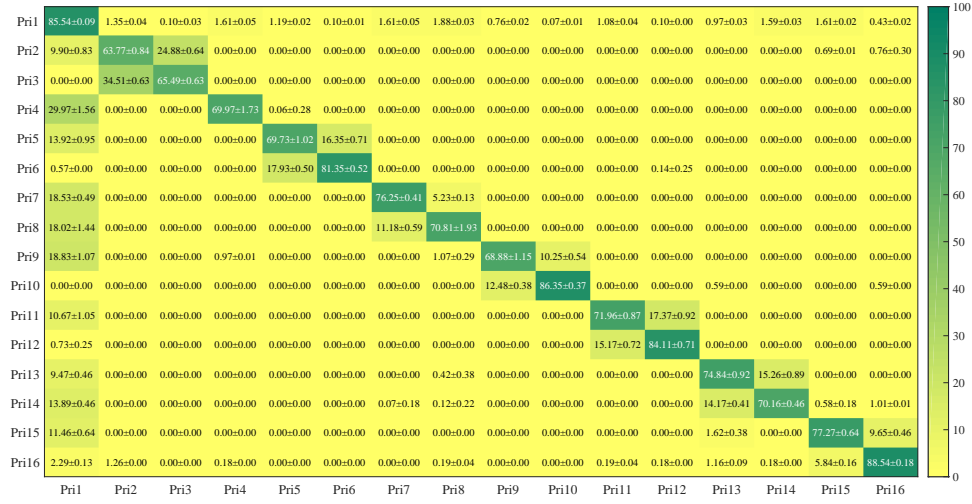
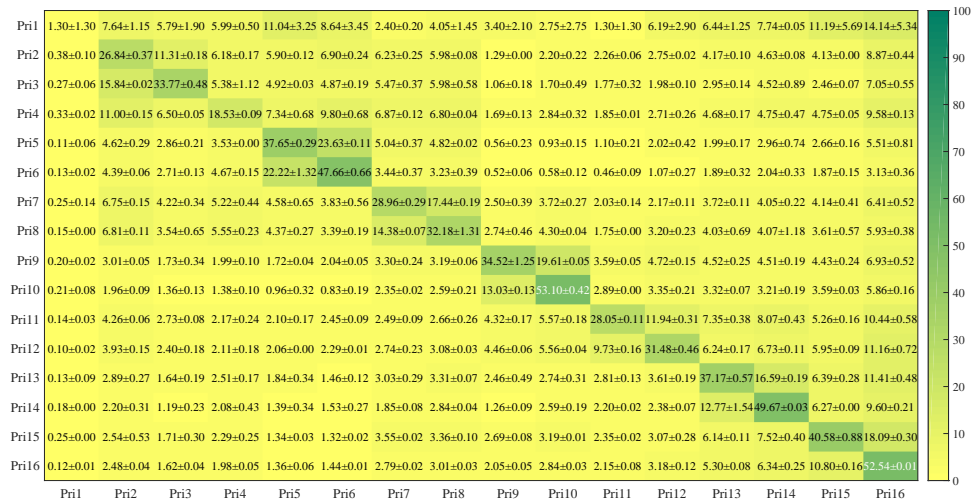


Figura 26. Matriz de confusión Movimientos primitivos con EMGs.



11.2. Matrices de Confusión de las actividades físicas

11.2.1. Resultados experimentales con la metodología de HMM_CyH

Figura 27. Matriz de confusión de actividades físicas con Kinect One®.

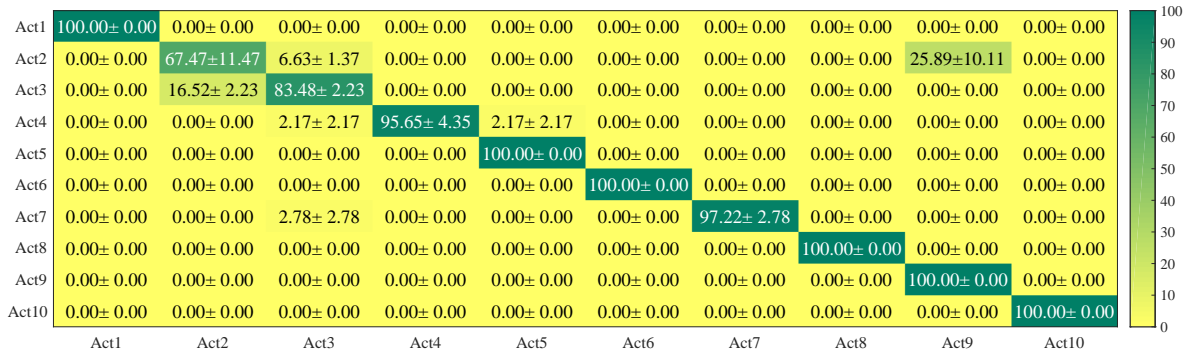


Figura 28. Matriz de confusión de actividades físicas con IMUs.

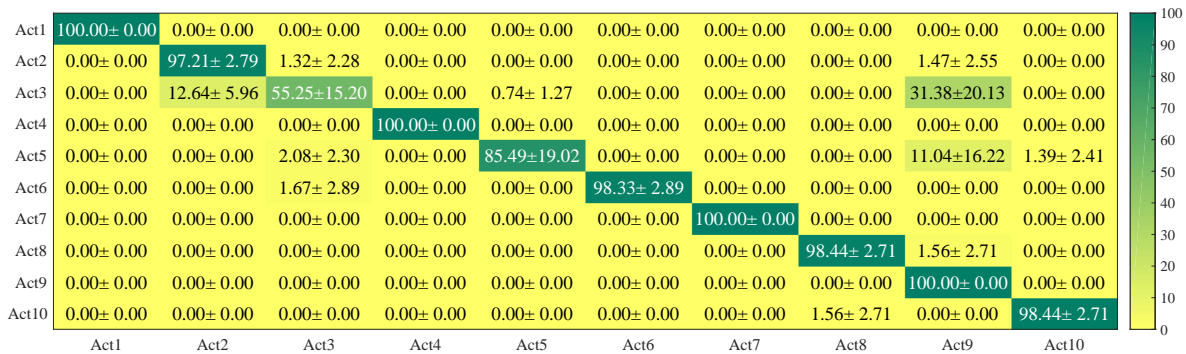


Figura 29. Matriz de confusión de actividades físicas con EMGs.

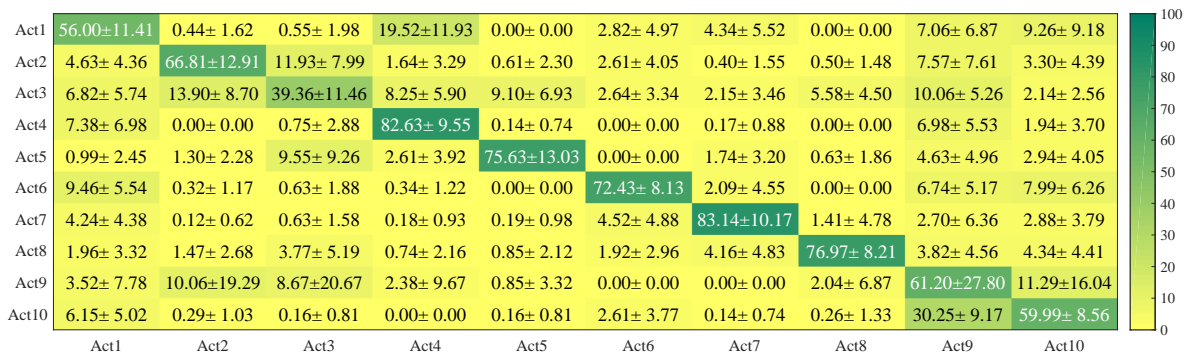


Figura 30. Matriz de confusión de actividades físicas con Kinect One®+IMUs.

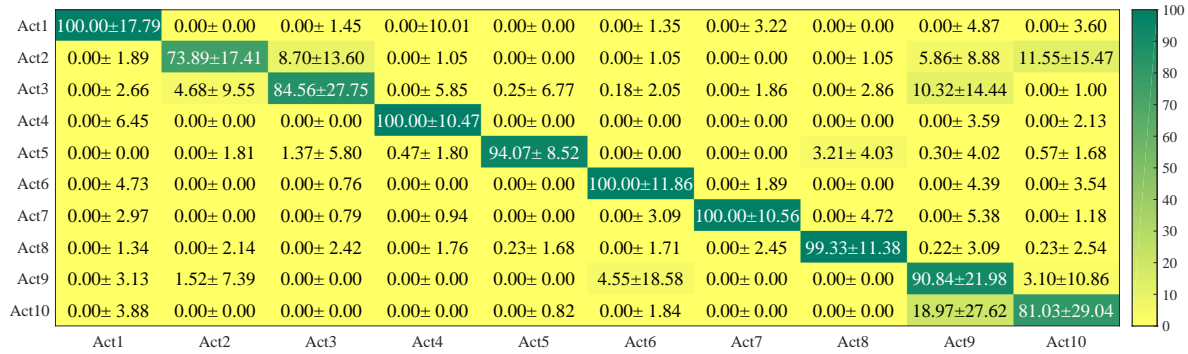


Figura 31. Matriz de confusión de actividades físicas con Kinect One®+EMGs.

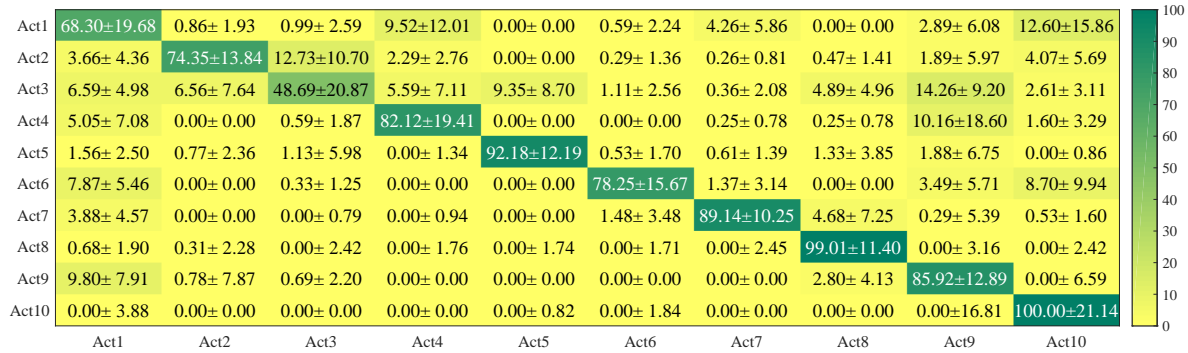


Figura 32. Matriz de confusión de actividades físicas con IMUs+EMGs.

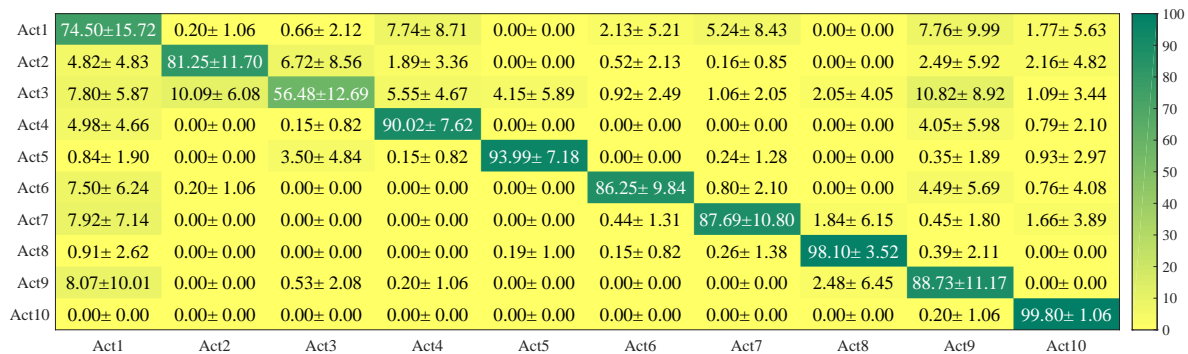
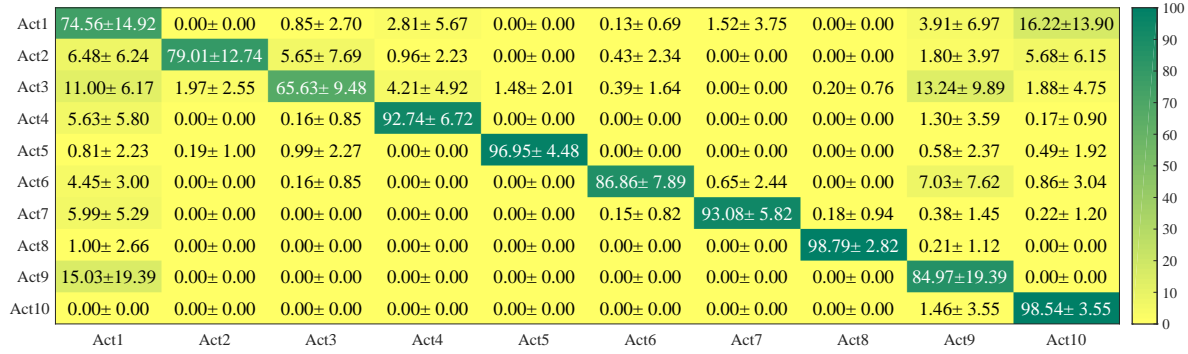


Figura 33. Matriz de confusión de actividades físicas Kinect One®+IMUs+EMGs.



11.2.2. Resultados experimentales con la metodología de HMM_C

Figura 34. Matriz de confusión de actividades físicas con Kinect One®.



Figura 35. Matriz de confusión de actividades físicas con IMUs.

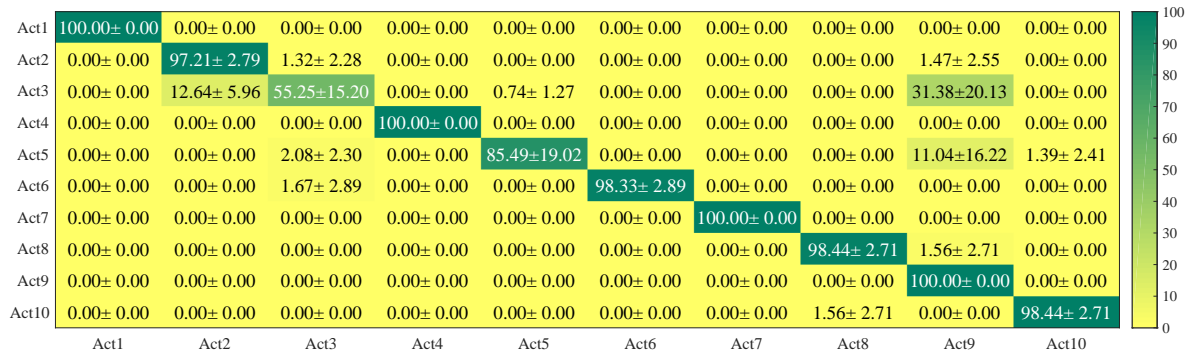


Figura 36. Matriz de confusión de actividades físicas con EMGs.

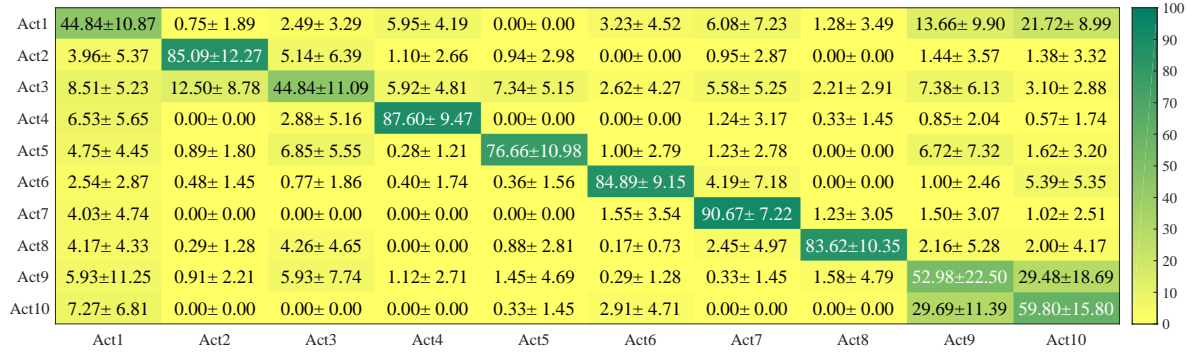


Figura 37. Matriz de confusión de actividades físicas con Kinect One®+IMUs.

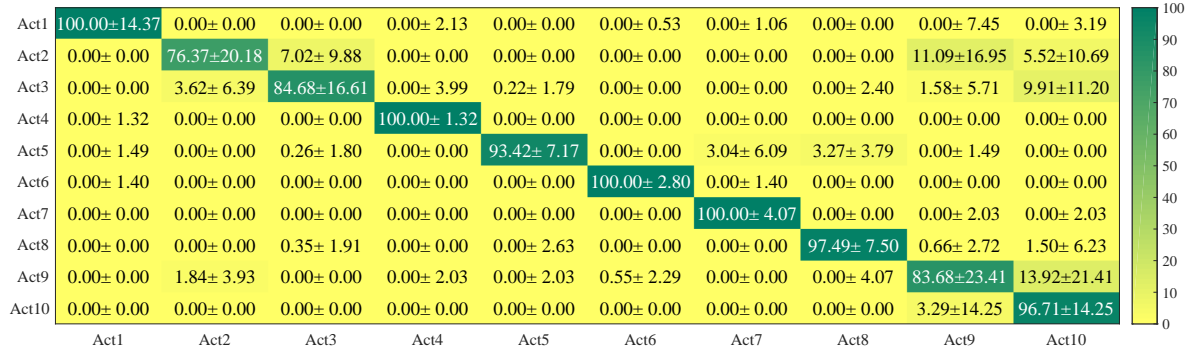


Figura 38. Matriz de confusión de actividades físicas con Kinect One®+EMGs.

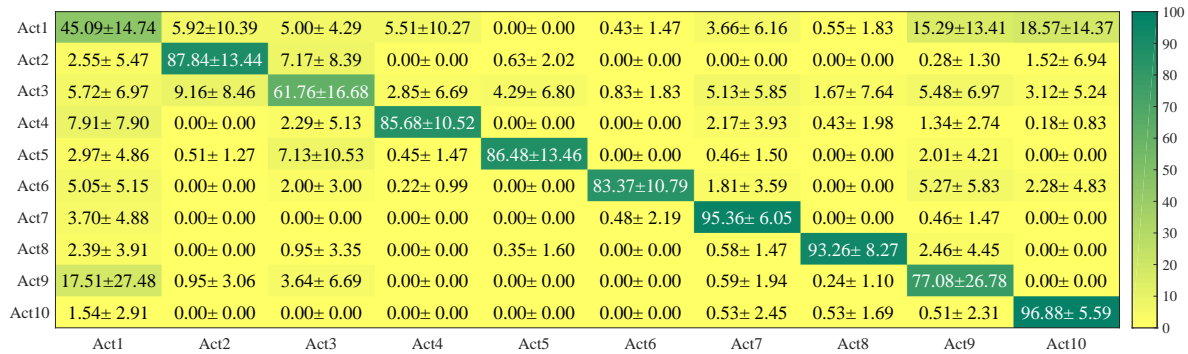


Figura 39. Matriz de confusión de actividades físicas con IMUs+EMGs.

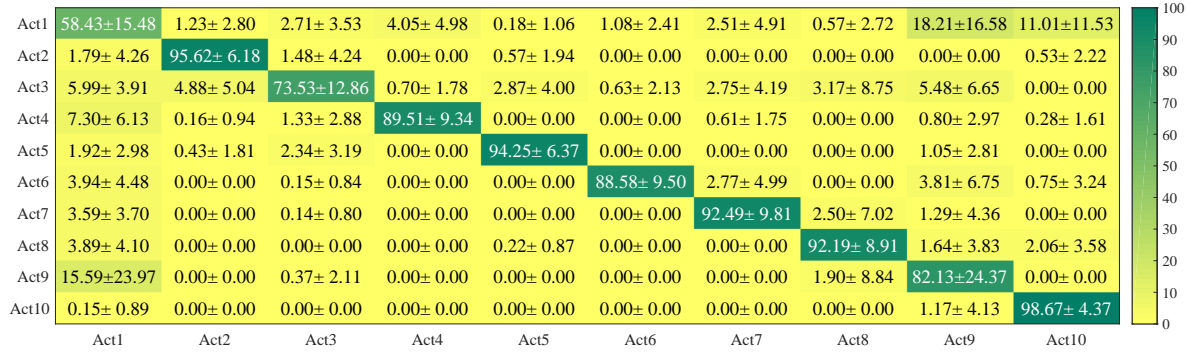
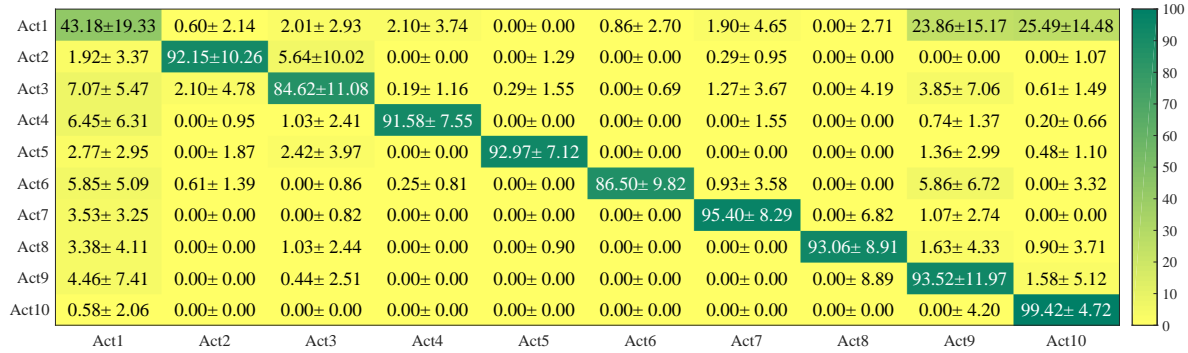


Figura 40. Matriz de confusión de actividades físicas Kinect One®+IMUs+EMGs.



11.2.3. Resultados experimentales con la metodología de HMM_BP

Figura 41. Matriz de confusión de actividades físicas con Kinect One®.

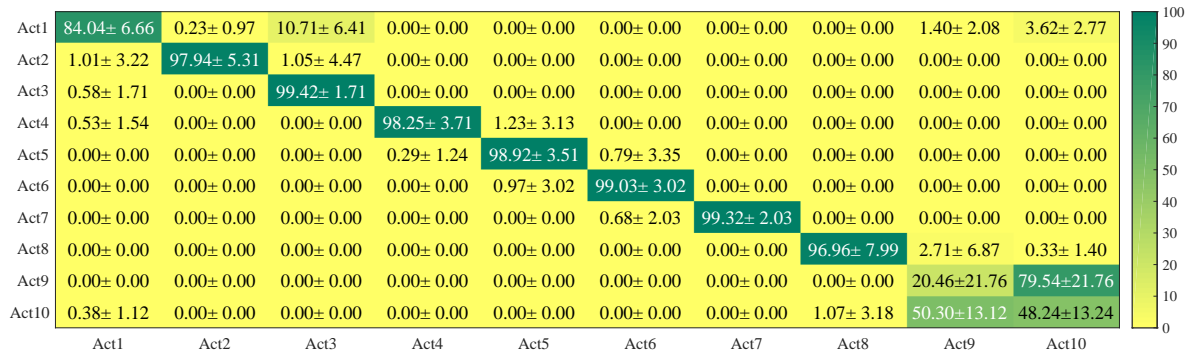


Figura 42. Matriz de confusión de actividades físicas con IMUs.

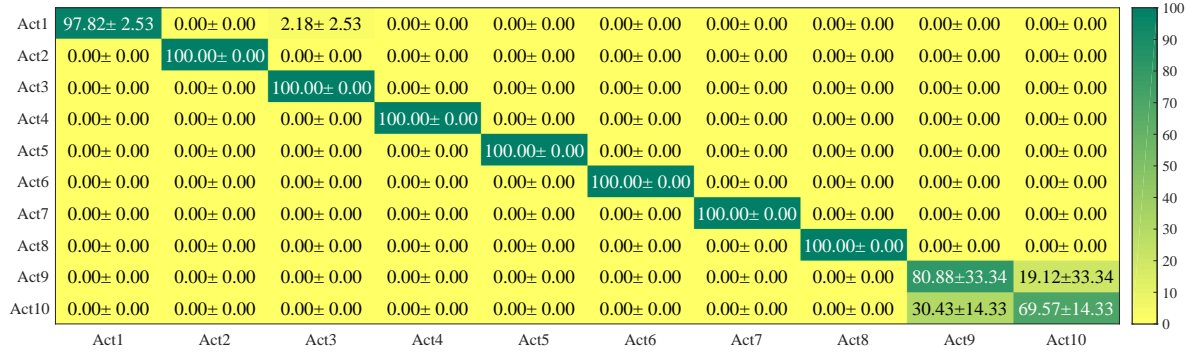


Figura 43. Matriz de confusión de actividades físicas con EMGs.

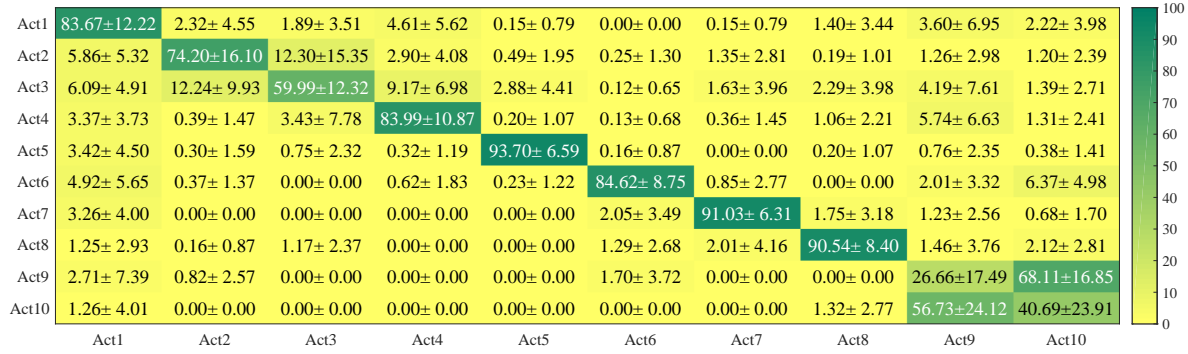


Figura 44. Matriz de confusión de actividades físicas con Kinect One®+IMUs.

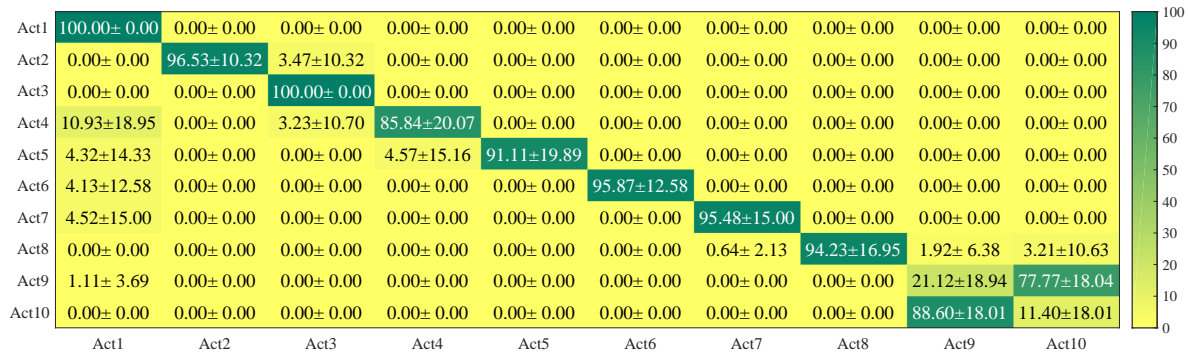


Figura 45. Matriz de confusión de actividades físicas con Kinect One®+EMGs.

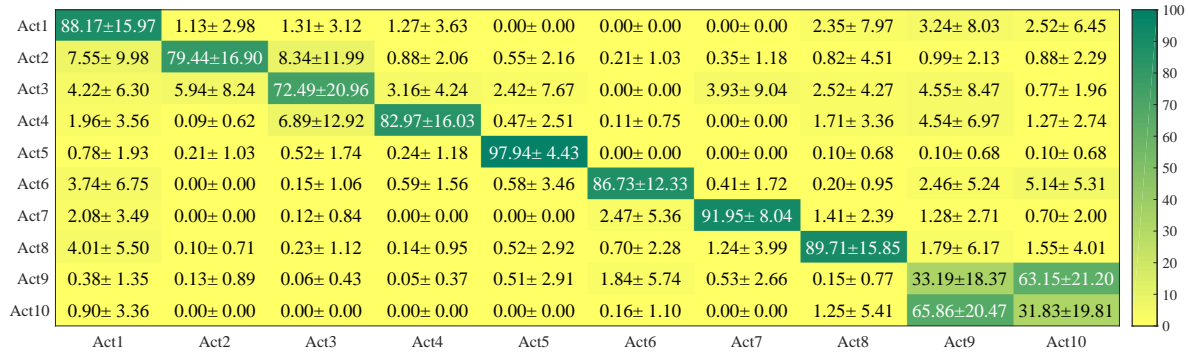


Figura 46. Matriz de confusión de actividades físicas con IMUs+EMGs.

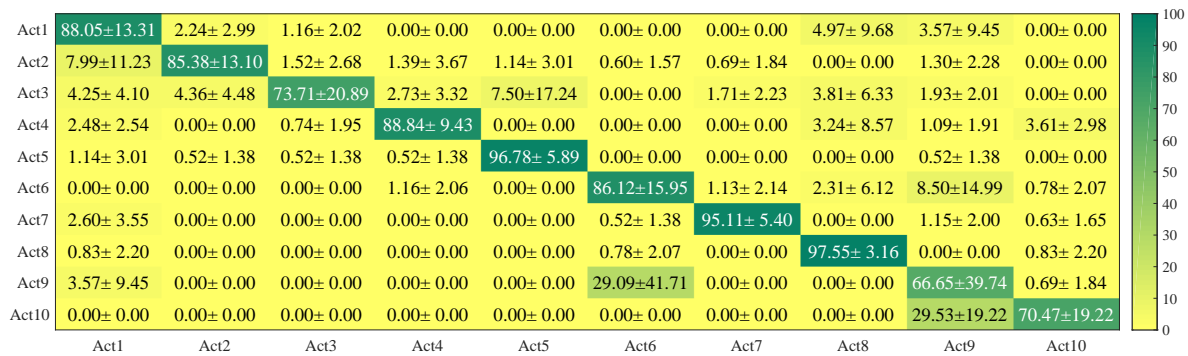


Figura 47. Matriz de confusión de actividades físicas Kinect One®+IMUs+EMGs.

